DOCUMENT RESUME

ED 150 842 FL 009 240

AUTHOR Garnica, Olga Kaunoff
TITLE Some Prosodic Characteristics of Speech to Young Children. Papers in Psycholinguistics and Sociolinguistics. Working Papers in Linguistics, No. 22.
INSTITUTION Ohio State Univ., Columbus. Dept. of Linguistics.
PUB DATE Feb 77
NOTE 63p.; Revised version of a Ph.D. dissertation, Stanford University, 1974 ; Some parts may not reproduce clearly due to print quality

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Child Language; Interaction Process Analysis; *Language Development; Language Patterns; *Language Research; *Language Styles; Language Variation; Paralinguistics; *Parent Child Relationship; Psycholinguistics; Speech Communication; Speech Habits; *Suprasegmentals
IDENTIFIERS Baby Talk; *Speech Registers

ABSTRACT

This study investigated the linguistic characteristics of speech addressed to the child and the features of the verbal environment critical for learning language. The study focused on the prosodic and paralinguistic features of adult speech to the young child. Adult speech directed to children was compared to other kinds of systematic speech adjustments adults make in response to differing social situations. Speech samples of 24 mothers addressing their children were analyzed. The mean age of one half of the group of 24 children was two years, three months; the mean age of the other half was five years, four months. The adult speaker made a systematic adjustment in the prosodic and paralinguistic aspects of speech addressed to the young child, and these adjustments varied in relation to the age of the child listener. It is suggested that various features of the baby talk register can serve at least two functions, analytic and social. (Author/SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED150842

SOME PROSODIC CHARACTERISTICS OF SPEECH TO YOUNG CHILDREN

A DISSERTATION

SUBMITTED TO THE COMMITTEE ON LINGUISTICS

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

BEST COPY AVAILABLE

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

By

Olga Kaunoff Garnica

November 1974

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

OSU Dept. of Linguistics

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

FL009240

c 1974

By

Olga Kaunoff Garnica

Preface

I wish to gratefully acknowledge the assistance of the following persons whose advice and support at various points in the progress of this work helped to bring it to a fruitful completion:

--Eve V. Clark, Charles A. Ferguson, and Ilse Lehiste, my principal advisors, for their expertise and encouragement;

--Dorothy Huntington and Clara Bush, who were extremely helpful members of my oral examination committee;

--John Ohala, director of the Phonetics Laboratory at the University of California, Berkeley for the use of the facilities and consultant personnel;

--the director of the language laboratory at Stanford University for the use of tape recording facilities and indispensible technical assistance;

--the mothers and children who participated in the study;

--many of my fellow graduate students at Stanford University who were available to exchange ideas and provide encouragement.

--finally, to my family and very especially to my husband for his immeasurable help and encouragement especially thru some of the more trying periods.

TABLE OF CONTENTS

Page

Preface 13

List of Tables and Figures 16

CHAPTER

1. INTRODUCTION
   - 1.1. Introduction 17
   - 1.2. Speech registers 18
     - 1.2.1. Definition of concept 18
     - 1.2.2. "Simplified" registers 20
   - 1.3. Baby talk register 23
     - 1.3.1. Types of studies 23
     - 1.3.2. Phonological and lexical features . . 24
     - 1.3.3. Syntactic complexity and redundancy features 25
     - 1.3.4. Prosodic features 27
   - 1.4. Hypotheses 29

2. METHOD
   - 2.1. Subjects 31
   - 2.2. Procedure 31
     - 2.2.1. Testing sessions 31
     - 2.2.2. Verbal tasks 32
       - 2.2.2.1. Picture task 32
       - 2.2.2.2. Puzzle task 34
       - 2.2.2.3. Story reading task 35
     - 2.2.3. Interviews 36
     - 2.2.4. Data collection. 37
       - 2.2.4.1. Taping of sessions 37
       - 2.2.4.2. Physical setting 37
       - 2.2.4.3. Instrumentation 38
     - 2.2.5. Data processing 38
       - 2.2.5.1. Selection of utterances . . 38
       - 2.2.5.2. Acoustic analysis 39
       - 2.2.5.3. Perceptual analysis 40
       - 2.2.5.4. Measurement of fundamental frequency and duration . . . 40

3. RESULTS
   - 3.1. Average fundamental frequency 42
   - 3.2. Frequency range 45
   - 3.3. Falling vs. rising sentence final pitch terminals 49
   - 3.4. Use of whispering 50

Page

3.5. Duration 51
3.6. Distribution of primary stress 54
3.7. Summary of results 55

4. DISCUSSION
4.1. Functions of characteristics of the baby-talk register 56
4.1.1. Prosodic features: Social function . . 57
4.1.2. Prosodic features: Analytic function . 58
4.2. Age of child listener 60
4.3. Speaker interviews 61
4.4. Implications for further research 64

5. CONCLUSION
5.1. Conclusion 66
5.2. Implications for theory of language development 66
5.3. Implications for linguistic theory 67

REFERENCES 69

List of Tables and Figures

TABLE Page

1. Age and sex of children participating with their mothers in the study 31
2. Description of materials in picture task 33
3. Comparison of fundamental frequency and frequency range data for C1 and C2 subjects 42
4. Average fundamental frequency for C1 and C2 subjects by session 42
5. Frequency range for C1 subjects by type of listener . . 46
6. Frequency range for C2 subjects by type of listener . . 46
7. Number of sentences with rising pitch terminals in the child-listener session by subject group and task 50
8. Average duration (in sec) of verbs in Puzzle Task . . . 52
9. Average duration (in sec) of color terms in Puzzle Task. 52
10. Number of instances of two primary stresses per sentence in six sentences from the Puzzle Task by type of listener 54
11. Presence and absence of some prosodic characteristics in adult speech by function and age of child listener . . 60

FIGURE

1. Mean fundamental frequency level for C2 subjects by type of listener 43
2. Mean fundamental frequency level for C1 subjects by type of listener 44
3. Mean fundamental frequency level by subject group and type of listener 45
4. Comparison of frequency ranges (in semitones) for C1 subjects by type of listener 47
5. Comparison of frequency ranges (in semitones) for C2 subjects by type of listener 47
6. Frequency range for C1 subjects by type of listener . . 48
7. Frequency range for C2 subjects by type of listener . . 49

CHAPTER 1

1.1. Introduction

It is now often claimed that speech addressed to children and speech addressed to adults differ in systematic and identifiable ways. This view represents a reversal of the direction taken by Chomsky (1967) and others (McNeill, 1970) who held the view that adult speech was mostly ungrammatical, replete with false starts, hesitations and slips of the tongue regardless of the addressee. This was supported to a certain degree empirically by Bever, Fodor, and Weksel (1965) who made calculations of grammatical and ungrammatical sentences in several speech samples.

This position had a strong influence on the theory of language development that was proposed. If the verbal input to the child is fragmented, confusing, and to a great degree unsystematic, then acquisition probably occurs independently of the linguistic environment. The input to the child must play a minimal role in the acquisition of language. Some verbal input was necessary, of course. This was clear from cases of children totally deprived of the opportunity to hear language. But input was relegated to a secondary role. Much of language acquisition was attributed to the child's innate capacities. The reasoning was that, if the language the child hears or overhears is not of a form that reveals the underlying systematicity of language, then some other mechanism must be available to the child in order for him to be able to abstract rules and gain knowledge about this systematicity.

Not all persons held this view. One of the strongest opponents was Labov (1970). He pointed out:

> The ungrammaticality of everyday speech appears to be a myth with no basis in actual fact. In the various grammatical studies that we have conducted, the great majority of utterances--about 75%--are well formed sentences by any criterion. When rules of ellipsis are applied and certain universal editing rules to take care of stammering and false starts, the proportion of truly ungrammatical and ill-formed sentences falls to less than two percent. (p. 42).

Although Labov makes no specific reference to speech directed to children, others had noted that the language heard by children is likewise neither phonologically nor grammatically deviant (Brown and Bellugi, 1964; Waterson, 1971).

What are the consequences of this alternative view of the role of the verbal environment for a theory of language development? One possible consequence is that greater emphasis is placed on the mother-child verbal interaction and the contribution of this experience to the child's learning of language. A direct result of this is that the importance of the innate mechanisms is deemphasized.

Another consequence is that the child's role becomes more important since the child can be seen as a more active participant in the language development process.

There are two questions relating to the verbal input to the child and his language development. First, what are the linguistic characteristics of speech addressed to the child? This is a purely descriptive question but necessarily preliminary to the second question: what features of the verbal environment are critical for learning language? Certainly, the latter question is more relevant to a theory of language development, and, more broadly, to the problem of cognitive development. Its answer, however, depends to some extent on the answer to the first question.

The present investigation seeks in part to answer the first question. It focuses on one linguistic aspect of adult speech to the young child, the prosodic and paralinguistic features of such speech, in an attempt to specify it more fully. In this sense, the study is intended to contribute at least indirectly to a more adequate theory of language development.

Finally, the findings of this study may contribute generally to a more adequate linguistic theory. Adult speech directed to children can be compared to other kinds of systematic speech adjustments adults make in response to different addressees in different social situations, i.e., speech registers (see 1.2.). Knowledge about the linguistic features characteristic of these adjustments and the distribution of their use will provide information which will have to be accounted for eventually by linguistic theory.

1.2. Speech Registers.

1.2.1. Definition of concept.

The term speech register refers to a type of social variation in language which is defined by the use to which language is put. Register is the most commonly used term, although the terms speech variety or speech style are also sometimes employed. All these are equally acceptable with perhaps the exception of the word "style" which has many strong associations with literature study and literary criticism that are not appropriate to the present use of the term. An excellent overview of the notion of speech register is available in Ellis and Ure (1969; and references therein). What follows is an elementary and highly schematic discussion of the basic points on the concept of register as preparation for the subsequent discussion.

If we assume that the focus of linguistic study is linguistic behavior, the reasonable object of study is language. In studying "language" we can be concerned with (a) different codes (tongues, languages), such as English, Chinese, etc., or (b) regional varieties within a single code, such as Boston English, Brooklyn English, Norfolk English, etc., or (c) social class varieties of a particular regional variant, such as the English of lower/middle/upper class Bostonians. The latter two are not mutually exclusive since social varieties are based to a degree on local variety. As Halliday,

MacIntosh and Strevens (1964) suggest:

> As one moves along the socio-economic scale, dialectal variety according to region diminishes. Finally at the apex there is no regional differentiation at all, except perhaps for the delicate shades which separate Cambridge and Oxford from each other and from the rest.

Registers are varieties also, but of a different sort than local and social varieties. They are set apart from them by virtue of the immediate circumstances of their use, i.e., the identity of the addressee(s) and other participants in the discourse, the purpose of the communication, etc. The particular register a speaker uses depends on the social situation that triggers its use. Changes in the register used are evidenced in changes in lexicon and in the phonological, syntactic, prosodic and semantic features employed.

There are rules which allow individuals to make judgments of appropriateness of register use and determine when a switch in register is required. These are socially-constrained rules which speaker and addressee have internalized as part of their communicative competence. The exact form and content of these rules is as yet incompletely specified although some things are known (see Ellis and Ure, and section 1.2.2 and 1.3 below).

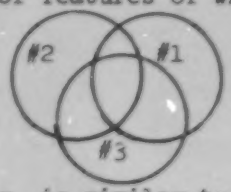
Universality. The notion of register seems to be a universal. Registers are known to operate in all speech communities which have been studied. The specific way in which a given register will operate in a particular speech community is determined by the specific social factors important in the community, how they function and the formal characteristics of the language.

Individual repertoires. Considered together registers cover the total range of language activity. They are not marginal aspects of language or "special" varieties. The entire range of registers reflects the range of social situations in a specific community or culture. They may be used (at least potentially) for specific purposes by all the speakers/writers of a language, thereby cutting across local and social dialectal varieties. Because having control of a register entails having knowledge of and experience with the social situation which governs it, no one individual controls the entire range of registers. No one person has experienced all the social situations in enough depth to learn the entire set. Thus, an individual's repertoire will contain more knowledge about registers in the social domains that he/she is involved directly and less in areas of less involvement. The particular repertoire an individual controls is governed by the common social factors of age, sex, education level, etc.

An individual's competence is more likely to contain rules for a wider range of registers receptively than productively. An adult is likely to control better the production of those registers necessary for functioning in the social domain where he is most active, while

on the receptive side having control over still other registers that he cannot or does not usually have to produce. For example, the male speakers or Berber in northern Africa have full knowledge of registers used productively exclusively by women, although they themselves would never use those registers in any circumstances (James Bynon, personal communication). The case of bilingual speakers is similar since they often have different distributions of register control in the two or three languages known to them. In many instances, this distribution is a function of the situation in which the language(s) was learned and the use(s) to which it is put.

Interrelationship of registers. Registers are interrelated in the sense that there is a great deal of overlap among them. They share many, in fact, most, of their features. One view is to consider the registers in a particular language (regional variety) as largely overlapping sets of features of which some portion of each set is not shared:



This view, however, is similar to the notion of isogloss in discussions of local varieties (dialectal studies) and shares many of the same criticisms. One such criticism is that the presence or absence of features is not absolute. It is more accurate to consider absence/presence as a continuum from greater to lesser frequency of occurrence. In this alternative view, the boundaries between registers are determined quantitatively and set apart from one another by the frequency with which a feature is used. Variables (features) are isolated and a register is defined in relation to other registers quantitatively.

Finally, registers cut across regional varieties to the extent that the social factors defining a particular register are present in the society. For example, it is hypothesized that all speech communities have a register appropriate for use to young children--the baby talk register. This is attested for many different kinds of languages (see 1.3 below) and seems plausible given that all speech communities have children who must acquire language and who must be socialized into the society. This is not to say that the distribution of individuals who control this register will be the same in all communities or that the features defining the register will be identical, but rather that one can isolate some variety in each of these speech communities that is appropriate for use to babies and young children.

1.2.2. "Simplified" speech registers

There is a subgroup of speech registers which have a sufficiently similar set of characteristics that they are sometimes grouped together in discussions on types of speech registers. Ferguson (1971)

has labelled these "simplified" registers. The term is used to refer to conventionalized varieties of speech which are utilized by speakers in situations where the addressee does not have full understanding of the language.

The label "simplified" indicates the process that is claimed to occur when the speaker is placed in a special kind of speech situation. When the level of linguistic competence between speaker and addressee is unequal, the speaker (here assumed to have full linguistic competence) is thought to adjust his linguistic output to a level that is "simpler," "more basic," "more "clarifying." What exactly can be called "simpler" is a matter open to question if one wants to establish strict criteria. However, on a more intuitive level many of the defining features of "simplified" registers do indeed appear to simplify, as will become more evident in the discussion of these characteristics. We know that it is possible for a scientist to take a complex theory, e.g., Einstein's theory of relativity, and "simplify" the content to its basics for the benefit of a less knowledgeable audience (say a freshman college student). In the same way it is possible to "simplify" an utterance in both content and form so that it would be suitable for use to a child with limited knowledge of language and the real world. The same or similar prepositional content would be conveyed but in a specific way, a "simplified" form.

It is likely that not all the "simplified" registers or varieties have been isolated or studied at this point, but some of the most commonly discussed examples include: (1) the baby talk register, the variety used primarily in addressing young children, (2) the foreigner talk register (Ferguson, 1971; Ferguson, 1972), (3) the foreign language classroom register (Henzl, 1973), and (4) the register used in speaking to partially deaf persons and other individuals who have physical (or mental) deficiencies that might affect language comprehension and use. At this point it might be useful to review some of the findings of the studies that have investigated these registers for the purpose of comparison and clarification. The information available is limited but sufficient for our purposes. The discussion here will focus on the foreigner talk register and the foreign language classroom register. The baby talk register is discussed in detail in section 1.3 below. The variety used in speaking to partially deaf persons and others has only been mentioned in passing. There has been no attempt to characterize it and therefore it will not be discussed.

The foreigner talk register. Foreigner talk has been studied by Ferguson (1971, 1962) experimentally and through investigation of published sources. The variety has defining phonological, syntactic, and lexical characteristics. Its phonological characteristics include: reduplication, the use of emphatic stress and intonation, slower, louder speech, and more distinct pronunciation. Grammatical characteristics include the omission, expansion and replacement of certain material. Omissions of articles, the copula, conjunctions, subject pronouns, and inflections marking the grammatical categories of case,

person, tense, and number in nouns and verbs are common. Expansions include realization of the subject of imperatives and the addition of tag questions. Replacements affect primarily negative particles, possessive pronouns, and subject pronouns. Lexical characteristics consist of substitutions and analytic paraphrasing, e.g. the use of one day gone and day before this for yesterday, and the use of special vocabulary, e.g. savvy for understand.

Although the characterization of this register is far from complete, this listing indicates the features which exist in speech directed to the addressee who has a poor command of the language of the speaker. It is not difficult to label this speech "simplified." Many of the characteristics serve to increase redundancy (e.g. reduplication) while others eliminate certain surface elements that are unnecessary and redundant for the purpose of expressing basic relationships between objects and events (e.g. omission of person markings on verbs that are redundant with the pronoun).

Foreign language classroom register. The foreign language classroom register also has defining phonological, syntactic and lexical characteristics (Henzl, 1973). The language used by speakers in this study was not English, but Czech. Some of the characteristics overlap with those defining the foreigner talk and baby talk register while others do not. The lack of sufficient data makes it impossible to decide whether these differences are due to the difference between registers or to differences between languages.

Some phonological characteristics of the foreign language classroom register are: few hesitation pauses and slips of the tongue, fully realized vowel quality, fully realized consonant clusters, and more pauses between sentences. Many of these characteristics are probably a direct result of the slower rate of speech that characterizes this register. The grammatical characteristics include the use of shorter sentences, fewer co- and sub-ordinate clauses, and a reduction in the use of inflection in nouns and verbs. In Czech the latter is expressed as an increase in the use of nominative and accusative cases (rather than instrumental) of nouns which in Czech have zero morpheme endings. The lexicon is characterized by the use of a more limited vocabulary, e.g. exclusively using plakat "to cry" rather than all the numerous variants plakala ~ řvala ~ bečela ~ breči, and more specific terms for less specific ones, e.g. "Mr. Smith" and "gentleman" used more than "he," "him".

The purposes of studying the simplified registers have been summarized by Ferguson (1972). He proposes that the study of these registers may clarify the notions of simplification in language and give us a better understanding of the factors that govern language use. In addition, the study of the foreigner talk register may be of value in understanding the process of pidginization (Ferguson, 1971). The study of the baby talk register may advance knowledge of the process of child language development (Ferguson, 1974; Slobin, 1967: 42-5). Thus, the study of these "simplified" registers has value above providing a taxonomic characterization of an aspect of verbal phenomena.

1.3. Baby talk register

1.3.1. Types of studies

The most extensively studied "simplified" register is the baby talk register. The reason for this is probably due to two major factors. For one, the interest in child language development has increased steadily since its beginning in the late fifties. One question which has been raised by many investigators is the effect of the language the child hears on his linguistic development. The second factor to play a role is the strong emphasis placed by some behaviorist theories on the role of language input in language development (Skinner 1957) and the consequent strong negative reaction to that view expressed so eloquently by Chomsky (1959). This ideological controversy brought the issue of the role of language input into the foreground but did not necessarily foster the study of speech directed to children acquiring language, since until the last five years the views of Chomsky prevailed to an overwhelming degree.

Studies of the baby talk register can be grouped into two types. They are sufficiently different in focus and methodology that they deserve to be discussed separately.

Type I studies. Studies of the first type were primarily carried out in the late fifties and throughout the sixties. The study of Arabic baby talk by Ferguson (1956) was one of the earliest studies of this type. The methodology of type I studies is remarkably homogeneous. The investigator gathers information on how adults talk to children in a particular speech community by eliciting from an informant. This information is often supplemented by observations from his own experience and also from information contained in published sources, e.g., novels, diaries, accounts, etc. The source of information is primarily one person, although some investigators consult other informants and occasionally observe a family. Information is usually limited to one or at most two regional varieties. No attempt is usually made to tape record speech.

Type I studies of the baby talk register have been conducted in fifteen languages. These include representatives from a wide range of language families of the world: Arabic (Ferguson, 1956), Berber (Bynon, 1968), Cocopa (Crawford, 1970), Comanche (Casagrande, 1948), English (Ferguson, 1964; Read 1962), Gilyak (Austerlitz, 1956), Greek (Drachman, 1973), Hidatsa (Voegelin and Robinett, 1954), Japanese (Fischer, 1970; Chew, 1969), Kannada (Bhat, 1967), Latvian (Ruke-Dravina, 1961), Romanian (Avran, 1967), Spanish (Ferguson, 1964) and Marathi (Kelkar, 1964). These studies have provided most of the information of a cross-language nature that is available.

Lexical and phonological features are those most commonly discussed in Type I studies, although some passing mention of syntactic devices is often included. The phonological features discussed usually exclude the topics of prosody except for occasional mention of pitch use and tempo (Kelkar, 1964 is an exception here). The focus is on the lexical items that are characteristic in the baby talk register, e.g. in English mommy and daddy for mother and father, and the phonological and morphological processes that distinguish these lexical items.

Type II studies. The studies of the second type came to the fore primarily in the early seventies, when there appeared a number of dissertations on the topic of mothers' speech to children learning language (Broen, 1972; Phillips 1970; Remick, 1971; Snow, 1972). Previous studies of parent-child speech input were conducted at Harvard (Brown, Cazden and Bellugi, 1969; Cazden 1965) and Berkeley (Slobin, 1969; Drach, 1969; Kobashigawa, 1969; Pfuderer 1969), but these were of a more limited nature, i.e. limited in variables studied or number of parent-child pairs observed. Other studies followed, some concentrating in greater detail on specific linguistic features. These are discussed further in 1.3.3. below.

To characterize Type II studies in a more general way, the source information on parent child speech in these studies are tape recordings of mother-child (in some cases non-mothers also) verbal interactions in a variety of structured situations, e.g. playing a game, telling a story, etc. The collection samples are transcribed and certain measures of grammatical complexity and redundancy are obtained. Statistical tests are employed extensively. Most of the measures are of a syntactic nature, although some attention is given to pauses and speech rate (esp. Broen) and most recently to semantic properties (Snow, 1974). All these studies deal with English speakers. No cross-linguistic comparisons are available.

In the two sections that follow (1.3.2 and 1.3.3) I will discuss the major findings of the Type I and Type II studies. Since these studies have been reviewed in detail elsewhere (Type I: Ferguson, 1964, 1974; Type II: Farwell, 1973; Vorster, 1974; Snow, 1974) the discussion will be brief. The discussion is divided into three parts:

(1) phonological features (other than prosodic) and lexical features of the baby talk register
(2) syntactic and redundancy features
(3) prosodic features.

The discussion of prosodic features leads directly into the purpose of the present investigation and the hypotheses to be studied.

1.3.2. Phonological and lexical features

Most of the information on the lexical features and the phonological features (semgental aspects) has been gained in Type I investigations of the baby talk register. Each of these studies has produced some 25 to 60 words which are claimed to make up the lexical items peculiar to the register. The major topics covered by these lexical items are:

(1) kin names and nicknames, e.g. daddy, mommy,
(2) body parts and bodily functions, e.g. footsie, night-night,
(3) basic qualities ("good", "bad", etc.), e.g. teenie, itty-bitty "little",
(4) names of animals and nursery games, e.g. doggie or bow-wow "dog", piggy-back.

These lexical items have certain foregrounding features. Reduplication is one common feature, as can be seen in items like night-night and bow-wow. Furthermore, diminutive suffixes appear on this set of lexical items in all the languages studied. The most common hypochoristic suffix in English appears in six of the example items above (doggie, teenie, etc.). In addition these lexical items have a similar canonical form, usually ending with an open syllable. The canonical form varies with the particular language, e.g. CVCV for Berber, (C)VCCV for Japanese, and CVC for Syrian Arabic. The last of these is an example of an exception to the generalization about final open syllables. Phonologically most of these "special" lexical items contain primarily stop consonants, nasals, and a limited selection of vowels. Of the twenty-seven lexical items listed for English by Ferguson (1964) only four contain sounds other than these. The sound "s" appears in footsie and pussy, and the sound "r" appears in burnie and birdie.

Certain phonological processes operate in the baby talk register. They include:

(1) the deletion or replacement of certain sounds in the adult phonological system, especially r sounds, e.g. English rabbit becomes wabbit (replacement), English drink becomes dink (deletion).
(2) the loss of certain phonological distinctions present in language directed to adults, e.g. Berber--the neutralization of vowel length distinction.
(3) distinct nasal assimilation, e.g. Gilyak damk "hard" becomes ama.
(4) replacement of velars by apicals, e.g. English tum on for come on.
(5) simplification of consonant cluster, e.g. English tummy for stomach.
(6) certain interchanges between sibilants, affricates and stops, e.g. Marathi--affricates replaced by stops.
(7) loss of unstressed syllables, e.g. Spanish tines for calcentines.

Many of these phonological processes move in the direction of less markedness and toward conforming with the phonological rules of children's speech at early stages of language development. This applies to the feature of reduplication as well. An example of the "less marked" rule is the replacement of r in rabbit with the semivowel w (wabbit). A counterexample to this rule, however, is the replacement of s by ch in Spanish, e.g. Spanish becho for beso.

1.3.3. Syntactic complexity and redundancy features.

A number of measures of syntactic complexity and redundancy show that speech to the young child is syntactically simpler and more redundant than speech to adults. This is manifested in a variety of ways. The following are the major characteristics of speech directed

to a child listener as opposed to an adult listener.

Co-ordination and subordination. Speech addressed to the child contains few or no embedded or conjoined clauses (Drach, 1969; Sachs, Brown, and Salerno, 1972; Snow, 1972; Phillips, 1973). The proportionate number of relative clauses, complements, and subordinate clauses is much lower in adult speech when the speaker is addressing a child. For example, Drach found that subordinate clauses are ten times more frequent in speech addressed to the adult than to the child.

Inflections. Speech addressed to a child listener contains fewer grammatical inflections than speech addressed to the adult. Snow (1972) found fewer inflections for both nouns and verbs. Many Type I studies also mention this as characteristic of other languages too.

Repetitions. There is a high incidence of repetitions in speech addressed to the child. One study (Kobashigawa, 1969) found that 35 percent of all utterances directed to the child listener are repetitions of some type (exact, partial, paraphrases). Not all sentence types are subject to repetitions equally. About 60 percent of imperatives, 25 percent of questions, and 15 percent of statements are repeated. Snow (1972) found that most repetitions are partial repetitions usually produced immediately after the full sentence form. Noun phrases and prepositional phrases were the parts of the sentence that were most frequently repeated. In addition, a number of utterances directed to the child listener are paraphrases. These may be viewed as a type of repetition.

Type/token ratio. The type/token ratio of speech directed to the child listener is smaller than that in speech directed to the adult (Broen, 1972; Drach, 1969; Phillips, 1970 and 1973; Remick, 1971). A small type/token ratio indicates a restricted and repetitious vocabulary. This finding and the immediately previous one point to the great amount of redundancy in speech addressed to children. This applies especially to children under the age of six years.

Number and length of sentences. Speakers use a greater number of sentences when addressing a child listener and these sentences are much shorter (Drach, 1969; Remick, 1971; Sachs, Brown, and Salerno, 1972; Snow, 1973). These two findings are undoubtedly related. Sentences used to the adult are an average of 2 1/2 times as long as those used to the child listener and about 10 percent of all utterances directed to the child consist of single word utterances (Broen, 1972). It is not surprising that more sentences are used to the child than to the adult in the same time period.

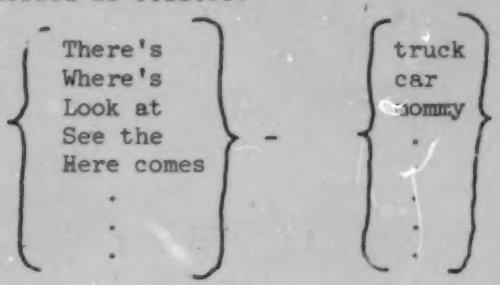
Verbal routines. One common verbal routine that introduces vocabulary items in a set frame has been noted by several investigators (Broen, 1972; Ferguson, Peizer, and Weeks, 1973). An example dialogue illustrates its use:

MOTHER TO CHILD: Look, Tommy.
Look at the truck.
See the truck.
Where's the truck.

CHILD: Truck.

MOTHER TO CHILD: Yes, truck.
Here comes the truck.

The underlined sentences indicate use of the set frame. In this example the frame sentences were produced with the same rhythm and heavy primary stress on the last word (truck). The frames may be schematized as follows:



Use of such frames has been found to be as common as five instances of the same frame produced within a five minute period (Broen, 1972).

Question and imperative sentence forms. The proportion of question and imperative sentence forms as opposed to declarative forms, increase in speech directed to the child listener. This feature is one of the best documented findings available (Blount, 1972; Broen, 1972; Drach, 1969; Pfuderer, 1969; Remick, 1971; Sachs, Brown, and Salerno, 1972; Snow, 1972). Questions, especially, are more frequent with estimates varying from 35 percent to 65 percent as compared to 70 percent declarative forms in speech to the adult listener. Imperatives account for about 30 percent of the utterances directed to the child. Ervin-Tripp (1970) suggests that a high degree of interrogative forms are the result of the greater need of adults to ask for feedback from the child.

Personal pronouns. The distribution of personal pronouns in speech to the child listener has been extensively investigated by Willis (1974). One of the most frequent findings in this area is the decrease in use of the third person pronouns in speech directed to the child (Broen, 1972; Snow, 1972; and a number of Type I studies).

The list presented here is not exhaustive but rather covers those findings that form the broad outlines of the syntactic complexity and redundancy features of mothers' speech to children.

1.3.4. Prosodic features.

Up to this point the discussion has focused on the particular lexical and syntactic characteristics of adult speech addressed to

young children. However, some investigators of the baby talk register have also noted the characteristic use of prosodic features, i.e. "features whose arrangement in contrastive patterns in the time dimension is not restricted to single segments" (Lehiste, 1970: 3). These features are referred to as pitch, stress, and quantity and may function on a paralinguistic level (Markel, 1965). Some acoustic correlates of these are duration, fundamental frequency (F_0) and intensity. These features play a role in all languages but in various ways. There are observations on the use of these features in the baby talk register for several languages.[1] The importance of such information for the student of language development is discussed in Crystal (1973). The focus of the discussion that follows is on the information presently available on the use of such features in the baby talk register by American English speakers.

Ferguson (1964) notes that even the casual observer may notice that adult speech to young children is characterized by higher overall pitch and a preference for certain intonational contours. Gleason (1973) also mentions a rise in the fundamental frequency of the voice when addressing the young child. This feature is again noted in Sachs, Brown, and Salerno (1972) in their study of speech to a two year old. They also note an increase in pitch change within a sentence and more instances of emphatic stress. Furthermore, they notice that in speech to the child the majority of sentences which they classified as interrogatives were, by their word order, simple declaratives with a rising intonation. Only a small portion of the "interrogative" sentences contained question words or had inverted word order. Sachs et al. suggest that rising sentential intonation may signal something other than "question" in speech to the child, and that the rising intonation may be a special kind of pitch change.

These same characteristics are mentioned by Grewel (1959) who astutely added a number of other features to the list on the basis of his own casual observations of adult speech to children. He notes a higher overall pitch and, in simple sentences, a rising intonation at the end of the sentence. He observes, further, that longer sentences are divided up in sections with each section having its own completed rising or falling contour. Frequent successive repetitions of the same contour are also noted. Speech to the young child is slower in tempo with obvious (prolonged--OKG) pauses between words, word groups, and particularly between sentences.

Grewel makes some further comments which though suggestive, are ambiguously stated and difficult to interpret. He asserts that "the dynamic accent in speaking to babies is strikingly diminished as compared to speaking to adults" and also that "when a dynamic accent is used, it is as it were compensated by a prolongation of the stressed word." (Grewel, 1959: 196). As an example of the later he offers the sentence: "No, that we ca_ _'nt do!" The term dynamic accent is not defined and could be interpreted as meaning any of several different things. From the example above, it seems at least likely that Grewel is using the term "dynamic accent" to mean emphatic stress. In any case, the prolonged duration of stressed words also

seems to be a probable characteristic of speech directed to the child.

Remick (1971) reports the only empirical evidence on prosodic feature characteristics. She studied the speech of ten mothers to their children (ages 1;4 to 2;5). She calculated both median fundamental frequency and frequency range from narrow band spectrograms for a subsample of sentences from each subject. The spectrographic analysis was run on fourteen to seventeen utterances per subject from each of two distinct speech situations: (1) speech directed to an adult, and (2) speech directed to the child addressee. Her finding was that only the mothers of the youngest children used a higher median fundamental frequency and a greater frequency range when addressing the child. The speech of mothers whose children had begun to acquire language showed "a dramatic restriction in both median and range" (Remick, 1971: 32).

Several methodological inadequacies in this study lead us to question the data on fundamental frequency and range as well as the conclusions drawn from them. Some of these inadequacies are acknowledged by the investigator. First, since the recordings were made in the subjects' home, the quality of the recordings was in all likelihood poor. The choice of sentences for spectrographic analysis was thus biased toward those with a more favorable signal to noise ratio. The investigator reports that only a limited number of readings could be obtained even from the measurable sentences. Second, in reporting the findings Remick gives no account of the procedure used for making measurements. This leaves open the question of how certain decisions were made, decisions that could have profoundly affected the values obtained. Third, there was no attempt to match sentences measured from each of the two situations in terms of their composition. A number of investigations (Peterson and Barney, 1952; House and Fairbanks, 1953; Lehiste and Peterson, 1961) have shown that vowels have intrinsic pitch, i.e. there is a connection between vowel quality and the relative height of the average fundamental frequency associated with it: higher vowels have higher fundamental frequency. Therefore, in any attempt to compare the fundamental frequency in two or more situations or even across subjects within a single situation, it is necessary to obtain measurements on the same verbal material. If the phonetic composition of the samples varies greatly, the differences observed may only reflect a difference in the composition of the two samples. Finally, no statistical tests were run on the frequency data. Therefore, it is not clear that the observed differences between speech to the child and speech to the adult were significant.

One final observation about prosodic features in speech to young children is that the use of such characteristics seems to diminish or disappear in most contexts by the time that child addressee is four to five years of age. This has been noted in the case of higher fundamental frequency (Gleason, 1973; Grewel, 1959). The situation for the other characteristics is as yet unknown.

1.4. Hypotheses

Research on the use of prosodic features in the baby talk register is a broad area of study. Although passing comments have been made on

this topic (section 1.3.4 above), practically no systematic investigation has been carried out. The present study represents an initial inquiry into this area, and will therefore, be limited to the study of only a few selected aspects of the problem.

The hypotheses to be tested in this study are:

Hypothesis 1: The use of prosodic features in adult speech directed to young children differs systematically from the use of such features in speech to other adults. (That such differences are evidenced in the corresponding acoustic parameters is presupposed.)

Hypothesis 1a: The fundamental frequency of speech to the child is higher than that of speech to the adult.

Hypothesis 1b: The frequency range of speech to the child is greater than that of speech to the adult.

Hypothesis 1c: The duration of syllable nuclei of words receiving primary stress is greater in speech to the child than the adult.

Hypothesis 1d: The use of a rising pitch terminal in declarative sentences is more frequent in speech to the child than in speech to the adult.

Hypothesis 1e: In speech to the child the basic sentence unit is more frequently segmented into shorter sections than in speech to the adult.

Hypothesis 2: The degree and character of all the above differences (hypotheses 1a-1e) vary as a function of the relative age of the child addressee. Generally the older the child is, the closer the use of prosodic features will approach the pattern in speech to the adult.

All of the above hypotheses are held to apply under experimental conditions where the speech context remains constant and only the addressees are varied.

Footnote to Chapter 1

1. An extensive commentary on the particular use of prosodic features in the baby talk register in Marathi is found in Kelkar (1964). Passing references to one or two features are also found in reports on this register in other languages (see bibliography).

CHAPTER 2: METHOD

2.1. Subjects

The subjects were twenty-four women college graduates under thirty-five years of age residing in the predominantly white, upper middle class suburban community surrounding Stanford University. The subjects were native speakers of American English with minimal or no knowledge of a foreign language. Their speech was devoid of any discernible speech disfluency. All the women had lived in California for at least one year.

The women were contacted initially by telephone and asked to participate in the study. They were told that the study was concerned with the ability of children of different ages to pay attention in a set of common, everyday situations. Each subject agreed to participate in two sessions. The subject was told that in the first session she would be familiarized with a set of situations and that in a second session she would interact with her own children in these situations. None of the subjects seemed aware that their own behavior rather than the behavior of the children was of primary interest.

Twelve of the women had a child in the 1;10 - 2;6 age range (mean age - 2;3). The other twelve women had a child in the 5;1 - 5;7 range (mean age - 5;4). There were an equal number of male and female children in each age group. Table 1 lists the age and sex of the children that participated in the study.

Table 1
Age and sex of children participating with their mothers in the study.

Subject	Age of Child	Sex of Child	Subject	Age of Child	Sex of Child
C101	2;6	M	C201	5;3	M
C102	2;1	M	C202	5;1	M
C103	1;10	M	C203	5;6	M
C104	2;5	M	C204	5;2	M
C105	1;10	M	C205	5;6	M
C106	2;2	M	C206	5;2	M
C107	2;6	F	C207	5;5	F
C108	2;5	F	C208	5;3	F
C109	2;6	F	C209	5;2	F
C110	2;0	F	C210	5;4	F
C111	2;3	F	C211	5;7	F
C112	2;6	F	C212	5;7	F
$\bar{X}$ age = 2;3 Age range = 1;10 - 2;6			$\bar{X}$ age = 5;4 Age range = 5;1 - 5;7		

2.2. Procedure

2.2.1. Testing sessions

Each adult took part in two testing sessions, and performed three verbal tasks in each session. In the first session the subject

directed her speech to another adult (adult listener session). The adult listener in all cases was the investigator. In the second the subject directed her speech to her own child (child listener session). In a counterbalanced design the adult-child session should precede the adult-adult session for half the subjects. However, it was highly desirable to obtain speech samples in both sessions that were as natural as possible. The possibility of justifying an adult-adult session following the adult-child session without revealing the purpose of the experiment seemed remote. On the other hand, all subjects accepted the request for the adult-adult session before the adult-child session for the purpose of familiarizing them with the materials.

2.2.2. Verbal tasks

The subjects performed three tasks in each testing session: (1) a picture task, in which the subject told a story about the persons and events depicted in each of a series of pictures, (2) a story reading task, in which the subject read a short descriptive passage, (3) a puzzle task, in which the subject gave a series of instructions on how to solve a puzzle.

The order in which the tasks were performed by each adult was different in each testing session. The adults in each group (C1 and C2) were randomly assigned to one of the ordering sequences in the adult-listener session and to a different one in the child listener session. For example, if a subject in the adult-listener session performed the tasks in the order: reading task, puzzle task, picture task, then the order in the child-listener session was different, e.g. picture task, story reading task, puzzle task.

2.2.2.1. Picture Task

Five colored pictures depicting situations thought to be of interest to children were chosen from several magazines. The pictures showed a boy eating a hamburger, some boys dressed up like Indians sitting around a campfire, a family on a picnic, a family doing household chores, and a little girl and her mother baking. Each picture was mounted on a 9" x 11" piece of colored cardboard. A short declarative sentence related to the events depicted in the picture was written below each photograph. Detailed descriptions of each picture and the accompanying sentences are given in Table 2. The subjects were asked to make up a short story to go along with each picture. They were told to incorporate the exact wording of the sentence accompanying the picture into their story. The instructions for the adult-listener session and the child-listener session went as follows:

Table 2
Description of materials in Picture Task.

PICTURE #1
Description: A small boy is sitting on a porch step. In his lap is a plate with a large hamburger. A large shaggy dog is standing next to the boy. The dog is eating out of his food dish.
Sentence: They are both hungry.

PICTURE #2
Description: Three small boys each wearing an Indian headdress and wrapped up in a blanket are sitting outdoors around a campfire. In the kettle boiling over the fire is some soup.
Sentence: It is cold.

PICTURE #3
Description: A family is having a picnic outdoors. The mother, father, and little boy are sitting on a blanket spread out on a grassy area. There is food in front of them. It is a sunny day.
Sentence: They are glad that it didn't rain.

PICTURE #4
Description: A large family (mother, father, numerous children, grandfather) is standing in front of their house. Each member of the family is holding some object indicating that they are preparing to do some household chore (the girl is carrying a large basket of laundry, the father has a bucket in one hand, etc.).
Sentence: Everybody is doing his chores.

PICTURE #5
Description: A woman and a young girl are in the kitchen. They are both wearing aprons. The woman is helping the little girl place some dough into a baking pan.
Sentence: Next time the girl will do it herself.

Adult-adult session
"I am going to show you pictures of some familiar events. Look at the pictures one at a time and tell a short story about the people and events you see. Each of your stories should be abour four or five sentences long. One of these sentences should be the one below the picture. Please give this sentence exactly as it appears there."

Adult-child session
"Here are the pictures you saw last time. (Hand pictures to subject). Today I would like you to tell _____ (child's name) a short story about the people and events pictures in

them. Your stories should be about four or five sentences long. One of these sentences should be the one listed below the picture. Please give that sentence exactly as it appears."

After the instructions were given, the subject was given the opportunity to ask questions about the task. Then the subject began the task. Since the task was both simple and straightforward, few problems arose. Occasionally in the adult-listener session, a subject would either forget to use the sentence accompanying the picture or would rephrase the sentence. In these cases the investigator commented on this and stressed the importance of including that sentence in the story exactly as it appeared. The subject was then asked to repeat her story, making the correction.

2.2.2.2. Puzzle task

In the puzzle task the subjects were asked to give verbal instructions to the listener on how to solve a puzzle. The puzzle was a small wooden object in the shape of a barrel. Solving the puzzle involved taking the barrel to pieces. The barrel could be taken apart only by pushing and removing certain pieces in a specified order. The subjects were told that the barrel could be taken apart and were given the necessary instructions. They were to give these instructions one by one to the addressee so that he/she could dissassemble the puzzle. The pieces composing the barrel were color coded to facilitate this process. The instructions were as follows:

1. Push in the green square.
2. Take out the piece.
3. Push in the red piece.
4. Take out the piece.
5. Push in the blue piece.
6. Take out the orange piece.
7. Take out the purple piece.
8. Take out the brown pieces.

The five year old children, by and large, had no problem identifying the correct pieces and following the instructions. Some of the two-year olds, however, either did not recognize some of the color terms used in the instructions or confused them. In these cases the mother was instructed that she could assist the child only after each instruction was presented as it appeared in (1)-(8) above. The instructions to the subjects on this task were as follows:

Adult-adult session

"This object is a barrel. (Show object to subject). It is a puzzle and can be disassembled into a number of pieces. The colored pieces must be removed in a certain order. This is a list of the steps a person would have to follow in order

to solve the puzzle. (Give subject list of instructions.) Give me the instructions one by one as they appear on the list and I will take the barrel apart. That way you can see yourself how the puzzle works."

Adult-child session

"Here is the barrel puzzle you saw me work last time. (Give subject puzzle.) This time _____ (child's name) will be taking apart the puzzle. You are to give _____ (child's name) the instructions on this list exactly as they are written there, in the order in which they appear. (Give list of instructions to subject.) You can assist _____ (child's name) if he/she has difficulty following any instruction but please give the instructions first as it appears on the list."

On a few occasions the pieces of the puzzle were moved incorrectly by a child so that the disassembly of the puzzle could not be continued. In these cases the investigator who was observing the proceedings interjected some instructions to the mother so that the situation could be righted and the process could continue.

2.2.2.3. Story Reading Task

A short passage about rainbows accompanied by a picture depicting a rainbow and other items mentioned in the passage (sun, raindrops, etc.) was presented to the subject. The passage was the first paragraph of the Rainbow Passage (Fairbanks, 1940):

> When the sunlight strikes raindrops in the air, they act like a prism and form a rainbow. The rainbow is a division of white light into many beautiful colors. These take the shape of a long round arch with its path high above, and its two ends apparently beyond the horizon. There is, according to legend, a boiling pot of gold at one end. People look, but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow.

The subject's task was to read the passage out loud. In the adult-listener session the subject simply read the passage out loud in the presence of the investigator. In the adult-child session the subject read the passage to the child and could interact verbally with the child if she wished. This was included because pilot subjects had expressed the desire to have such an option. They felt that such interaction more closely approximated the manner in which they read stories to their children at home.

The instructions to the subjects were as follows:

Adult-adult session
"Accompanying this picture is a short story about rainbows. (Give picture to subject.) So that you will have read the passage through once, please read the passage out loud now."

Adult-child session
"This is the story about rainbows which you read last time. I would like you to read the story to _____ (child's name) as you might read a story to him/her at home. If you like, you may point out objects in the picture which are mentioned in the story or ask _____ (child's name) questions about them. If, while you are reading, you are interrupted in the middle of a sentence, please start again at the beginning of that sentence and proceed from there."

The Rainbow Passage was chosen for several reasons. First, the subject matter was thought to be of interest to children even though some of the sentences were quite complex. And, second, Horii (1972) has shown that there is a high correlation (+.98) between average fundamental frequency measurements for the second sentence and average fundamental frequency measurements for the rest of the passage. This suggests that measurements made in the second sentences would be very close to the values obtained if every sentence in the passage was measured. Thus, it would be possible to measure a small sample (one utterance per subject) and these measurements would be generalizable to the entire passage.

2.2.3. Interviews

An interview was conducted with each subject upon completion of the verbal tasks in the child-listener session. The interview was conducted in the presence of the child, who was occupied with some toys provided by the investigator. The purpose of the interview was to determine whether the subjects in the study were aware that they modified their speech when addressing their children, especially whether they noticed any prosodic modifications. If they were aware of such changes, what kinds of modifications did they notice?

The questions included in the interview appear below. The rather broad (imprecise) terms "same sort of voice" was used in an attempt to elicit comments from the subjects on the prosodic aspects of their speech. The probes were also structured to steer the subjects' responses in this direction. A direct question containing the term "prosodic" was ruled out because of the difficulty in explaining the meaning of this highly specialized terminology.

Interview Questions:

(1) Did you notice whether you spoke with the same sort of voice when addressing _____ (child's name) as you did when you addressed me in the previous session?
If "yes" answer: Was your voice different? What sort of differences did you notice? PROBE (if necessary):

Was it higher or lower in pitch? Softer or louder? Was there more or less fluctuation in the pitch of your voice?

(2) Have you noticed any such differences on other occasions or in other situations? On what occasions? In what circumstances?

(3) Have you ever observed that other parents speak in a different sort of voice to their children than they do to an adult? Who was it? What differences did you notice? On what occasions did this occur? What were the circumstances involved?

(4) If you want to get your child's attention, or get him/her involved in conversation, how would you do this? Specifically, would you change your voice in a particular way? What way(s)?

After the interview was completed, the subjects were informed of the actual purpose of the study and of the general hypotheses and predictions advanced by the investigator. They were encouraged to pose questions regarding any aspect of the study. All the subjects concurred that they had not been aware that their own speech rather than the child's behavior was the primary focus of the study.

2.2.4. Data Collection

2.2.4.1. Taping of sessions

All the testing sessions were recorded on tape in their entirety. The subjects were told that all the sessions were being recorded, including the adult sessions. The reason given for taping all the sessions was that the study had many participants and that the investigator would therefore be unable to rely on her memory alone in reviewing the various sessions. Most of the subjects accepted this reasoning without any further questioning. A few questioned the necessity of taping the adult-listener session when the child was not present. They were told that the investigator would review the adult-listener session prior to the child-listener session to determine if the tasks were performed in the manner expected. This was, in fact, done in a number of cases. No subject persisted in questioning the reasons for taping the testing sessions once these reasons were given.

2.2.4.2. Physical setting

All the testing sessions took place in an acoustically treated room, ordinarily used for recording by the university foreign language laboratory. Its dimensions were 6' by 9'. This room was selected in order to ensure tape recordings of a high quality. The room was on the inside of the building and had no windows. The only outlet was to a short hallway that led to the inside corridor of the building. The experimental room was equipped with special features to minimize the effect of sound waves reflecting off the flat surfaces--floor-to-ceiling buffers lined a portion of each wall, the floor was covered with a thick wall-to-wall carpet, and acoustical tile covered the

ceiling. The 2" thick solid wood door was insulated on all sides for a tight seal with the door frame. The room was illuminated by an overhead incadescent light to avoid the low frequency noise ("hum") often emitted by fluorescent bulbs.

The tape recorder was placed on a small table in one corner of the room. Next to the table was a chair where the investigator sat during the child-listening session. Two large, brightly-colored floor pillows were placed in the middle of the room parallel to one another and approximately 18" apart. During the testing sessions, the speaker (the subject) and the addressee (the investigator in the adult-listener session; the child in the child-listener session) each sat on one of the pillows facing one another.

The exact placement of the speaker and the microphone was determined by a preliminary test. The investigator read aloud a short excerpt from the Rainbow Passage in a normal voice several times. These readings were recorded in the experimental room on the same equipment used during the testing sessions. The speaker's position in the room was varied, as well as the left or right deviation from the straight ahead position of the speaker's head in relation to the microphone. Subsequently these recordings were judged by the investigator and one other person for quality of recording.[1] A minimum of distortion and fluctuation as a result of head movement was sought. On the basis of these sample recordings, the two judges determined the location of the speaker within the physical layout of the room that yielded the best recording. In the study the subject always sat in this location.

2.2.4.3. Instrumentation

The tape recordings were made on a Revox A77 tape recorder using a Sony Electret condenser microphone and Scotch 176 tape (1/4" x 1200'). The tape recorded was calibrated just prior to the beginning of the study to give a flat frequency response of ±2 db over the range of 50-10,000 Hz. The calibration was checked once approximately half-way through the study, and a second time upon completion of the study. The tape recorder performed reliably. The microphone was attached to a lavalliere placed around the subject's neck. The microphone hung approximately 10" from the subject's mouth.

2.2.5. Data Processing

2.2.5.1. Selection of utterances

A total speech sample of 30-40 minutes was obtained for each subject (both sessions combined). Since it is extremely costly in terms of time and data processing equipment to analyze such an enormous amount of data, a subsample of the utterances was chosen for perceptual and acoustical analysis. The utterances selected for analysis were the sentences provided in the picture task (5 sentences), the instruction sentences from the puzzle task (8 sentences), and the second sentence from the reading task (1 sentence). This yielded a sample of 14 sentences from each of 48 testing sessions. A total of 672 sentences were analyzed.

These particular sentences were selected in order to make inter-session and inter-subject comparisons on samples in which the lexical items were the same. With lexical content held constant, the analysis could focus on the properties that were of interest in the study.

2.5.2.2. Acoustic Analysis

Instrumentation. The utterances selected for analysis were dubbed from the original recordings using a duplicate Revox A77 tape recorder. The dubbing was done by means of a machine-to-machine patch cord leading from the output jack of machine No. 2. These dubbed utterances were then processed on the Pitch Extractor in the Phonetics Laboratory at the University of California at Berkeley (Krones, Ms.).[2] This Pitch Extractor produces a display indicating the fundamental frequency (F_0) of the voiced portions of utterances.

The Pitch Extractor is an analog device operating in real time so that the pitch contours may be recorded. Its output is a voltage that ranges from -10 to 0 volts which varies according to the frequency of the input signal. The Oscillomink is used to record this voltage and display the pitch contour on a roll of calibrated paper which is approximately 5 1/2" wide. The top half of the display is used to record the pitch contours while the bottom half is used to record amplitude (voltage). The Pitch Extractor was used in conjunction with a Transpitchmeter which supplied the input filter and calibration tones.

The Pitch Extractor can be adjusted for the frequency range that it will measure. The maximum frequency to be measured can be varied from 100, 150, 200, 300, 400, or 600 Hz. The lower limit to be measured is then represented by a percentage of the maximum frequency--25 percent, 50 percent, 75 percent. In the processing of utterances it is first necessary to select a range of measurement for each set of sentences. Once this range is determined, frequencies above the maximum are clipped off automatically and frequencies below the lower limit indicating lack of phonation, are not calibrated for frequency. A range must be set for each set of utterances that will encompass the range of frequencies represented in the voiced portions of the utterances. Calibration tones are used to indicate the frequency represented at particular points in the display within the established range.

Initially, a small subset of the utterances selected for analysis were processed on the Pitch Extractor. Twenty-four utterances from three speakers (2 C1 subjects and 1 C2 subject) were chosen for this preliminary analysis. There were an equal number of utterances from the adult listener session and the child listener session for each speaker. These utterances were selected as representative of the entire sample in terms of quality of recording and pitch range.

The purpose of this preliminary analysis was to determine whether the dubbed tape recordings produced satisfactory displays and whether it would be necessary to set different calibrations for the same subject when analyzing utterances from the adult-listener and child-listener

sessions. The resultant displays were judged to be satisfactory for measurement purposes. However, there were sufficient differences between the two sessions, as well as among subjects, to suggest the need for frequent recalibrations from session to session.

The total set of utterances was than processed on the Pitch Extractor. The machine was optimally calibrated for the utterances produced by each subject in each testing session. Calibration tones were used to record the frequencies represented on the displays for each new calibration. The displays were produced at what was judged to be an optimum rate--100 mm/sec. The displays produced for each subject were inspected for instances where the frequency was outside the maximum or minimum of the optimal frequency range. Utterances in which this had occurred were processed again with a new calibration.

2.2.5.3. Perceptual Analysis

The utterances selected for analysis were transcribed by the investigator. The total transcription consisted of a broad phonetic transcription of the segmental portion of each utterance in IPA notation, and a transcription of the accentual pattern. Four levels of stress were marked: Stress 1 (primary stress), Stress 2 (secondary stress), Stress 3 (tertiary stress), and Stress 4 (unstressed).

2.2.5.4. Measurement of fundamental frequency and duration

Measurements of fundamental frequency and duration were made for each syllable nucleus in each utterance. The following information was recorded: (a) the fundamental frequency at the beginning, peak, and end point of the syllable nucleus, (b) the location of the peak within the syllable nucleus, (c) the duration of the syllable nucleus and (d) the intensity at the peak of the syllable nucleus.

Clear, plastic templates were constructed from the calibration sheets for measuring fundamental frequency. A separate template was made for each calibration. Horizontal lines representing calibration tones were drawn parallel to the base line appearing in both the calibrations and the frequency display recordings. The lines represented specific frequencies on the Hertz (Hz) scale, e.g. 210 Hz, 220 Hz, etc. The interval between each pair of lines was 10 Hz. The template was superimposed on the fundamental frequency display for each syllable nucleus and the frequency of the beginning, peak, and end point was determined. If the point occurred exactly in the middle of an interval, it was assigned a value half-way between the two values defining the interval. For example, a point in the middle of the interval between 210 Hz and 220 Hz was assigned the value 215 Hz. Points which occurred within the interval but which were closer to one of the two values were assigned that value. Measurement error was estimated at ±5 Hz.

In most cases the syllable nucleus corresponded to a separate syllable in the utterance. However, in certain cases the boundaries between two syllables were not well enough defined on the frequency display and the two syllables were considered, for measurement purposes, as a single syllable nucleus. For example, this was the case for the

words in the in some speakers' renditions of the sentence "Push in the [ɪnə] green square."

Occasionally the end point of the last syllable nucleus in an utterance was impossible to measure because the subject's voice exceeded the limitations of instrumentation at the lower values. It was expected that such instances of unmeasurable phonation would arise. The frequency with which this occurred varied with the subject and type of session. In no case did this occur, in more than 5 percent of the sentences in one session. When it did occur, the lowest observed frequency value for that subject was assigned.

Duration was measured from the beginning point to the end point of each syllable nucleus. A transparent metric ruler was used. Since the display was produced at the rate of 100 mm/sec, one mm on the ruler was equal to .01 seconds. The location of the frequency peak in the syllable nucleus was also recorded.

The intensity curves for each utterance were recorded directly below the frequency display. Intensity was not calibrated to an absolute scale. Therefore, the values for intensity corresponding to the peak of a syllable nucleus were recorded in millimeters. This made it possible to compare the intensity at the peak of one syllable nucleus with the intensity at the peak of another syllable nucleus.

Intra-observer reliability was obtained for measurements of fundamental frequency and duration. A 10 percent sample of sentences was chosen randomly for remeasurement. Remeasurement of duration yielded values identical to the initial measurement. Therefore, no statistical analysis to determine reliability was performed. Remeasurement of fundamental frequency did yield slightly different results. The correlation between the initial fundamental frequency measurement and the remeasurement values was +.97. The intra-observer reliability was considered acceptable.

Footnote to Chapter 2.

1. I am grateful to Dr. Dorothy Huntington for giving her expert advice on this matter.

2. I thank Jean Marie Hombert for his assistance in processing this material.

CHAPTER 3: RESULTS

3.1. Fundamental frequency

The average fundamental frequency was computed for each subject for each session from the speech samples that were measured. These data are shown in Tables 3 and 4.

Table 3
Comparison of fundamental frequency and frequency range data for C1 and C2 subjects.

	C1 subjects (N=12)		C2 subjects (N=12)	
	Adult listener	Child listener	Adult listener	Child listener
Mean fundamental frequency (Hz)	197.6	267.3	202.8	206.4
Mean fundamental frequency (Hz)	43.2	48.4	43.6	43.9
Total range (st)	10.5	19.2	10.9	12.6

Hz = hertz
st = semitones above the zero frequency level of 16.35 Hz.

Table 4
Average fundamental frequency for C1 and C2 subjects by session.

	C1 subjects					C2 subjects			
Subject	Adult listener (Hz)	(st)	Child listener (Hz)	(st)	Subject	Adult listener (Hz)	(st)	Child listener (Hz)	(st)
C101	201	43.4	238	46.4	C201	185	42.0	193	42.7
C102	181	41.6	240	46.5	C202	202	43.5	208	44.0
C103	188	42.3	274	48.8	C203	215	44.6	216	44.7
C104	178	41.3	211	44.5	C204	188	42.3	197	43.1
C105	195	42.9	269	48.5	C205	199	43.3	198	43.2
C106	215	44.6	300	50.4	C206	239	46.4	241	46.6
C107	189	42.4	276	48.9	C207	198	43.2	197	43.1
C108	241	46.6	302	50.5	C208	186	42.1	188	42.3
C109	178	41.3	272	48.7	C209	206	43.9	205	43.8
C110	193	42.7	296	50.1	C210	206	43.9	223	45.3
C111	207	43.9	274	48.8	C211	234	46.1	237	46.3
C112	205	43.8	256	47.6	C212	176	41.2	174	40.9

Hz = hertz

st = semitones above the zero frequency level of 16.35 Hz

The mean fundamental frequency for the C1 group was 197.6 Hz in the adult-listener session and for the C2 group it was 202.8 Hz. These figures are within the expected range of values for female speakers (Linke, 1953; Peterson and Barney, 1952; Snidecor, 1951). The variations from speaker to speaker result primarily from differences in the size of the vocal folds. Other factors that can affect the values obtained are the nature of the verbal material spoken by the subject and the particular measure that is used to compute the values (mean versus median). The difference between the means for C1 and C2 subjects in speaking to an adult was not significant ($t = .75$; $df = 11$; $t_{(.05)} = 2.20$).

The mean fundamental frequency for the C1 and C2 subjects in the child-listener sessions were 267.3 Hz and 206.4 Hz respectively. For the C2 subjects, where the listener was a 5 year old, the frequency level in speech to the child was not very different from the level in speech to the adult. The difference between these two means is not significant ($t=2.07$; $df=11$; $t_{(.05)}=2.20$). The small differences between the two types of sessions is evident in the graphic representation in Figure 1. In only one C2 subject was the difference between speech to the adult and speech to the child more than one semitone (1.4 sts).

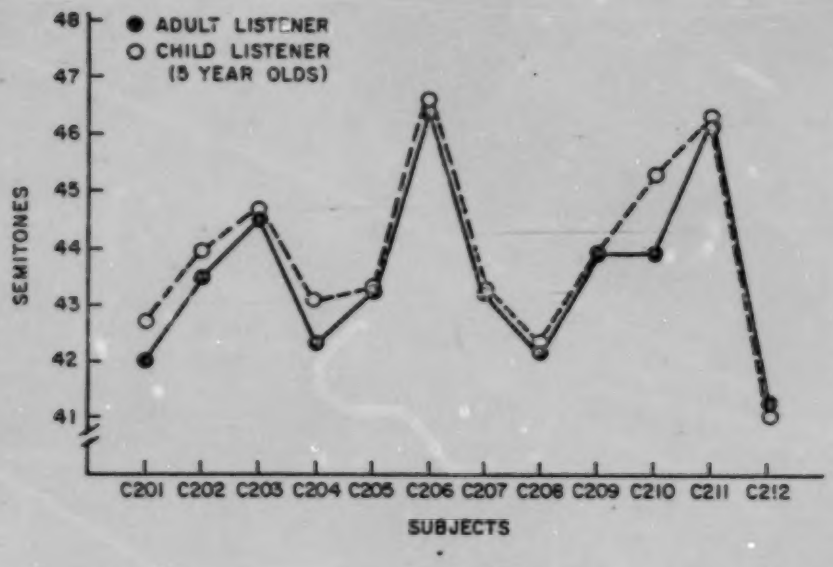


Figure 1. Mean fundamental frequency level for C2 subjects by type of listener.

The difference between the means for the child listener and adult listener conditions for the C1 subjects is quite large--197.6 Hz versus 267.3 Hz. This difference is highly significant ($t=11.55$; $df=11$; $t_{(.001)}=3.11$). For all C1 subjects, the fundamental frequency

level in speech to the child was considerably higher than in speech to the adult. The differences between the two types of sessions ranged from 3.0 semitones up to 7.4 semitones. These large differences are quite evident by visual inspection of Figure 2.

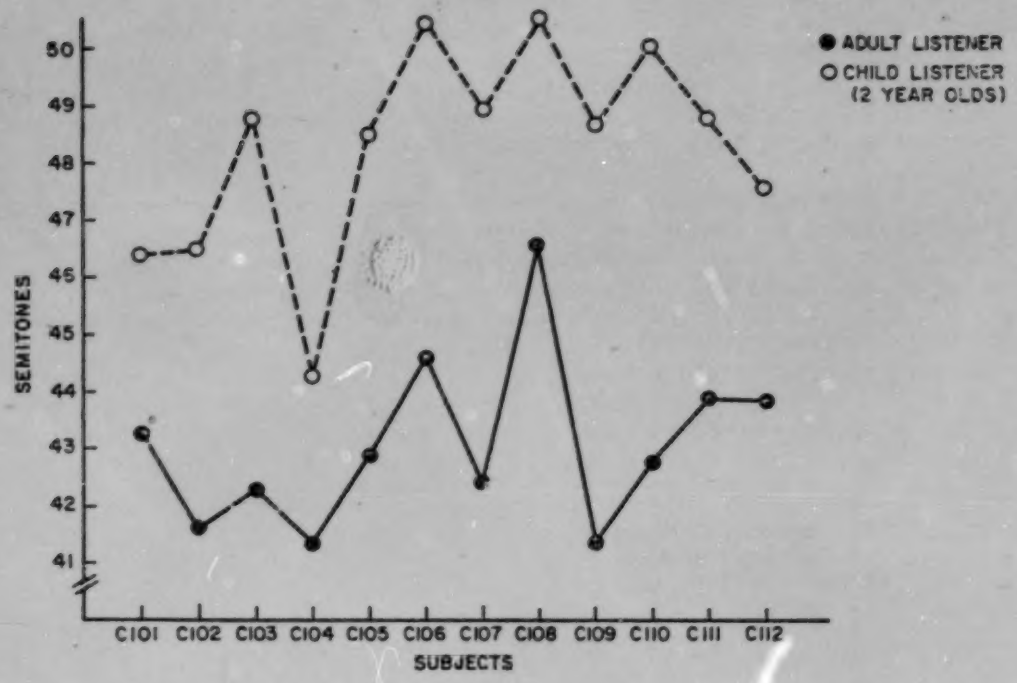


Figure 2. Mean fundamental frequency level for C1 subjects by type of listener.

It is evident that the subjects used a higher pitched voice when speaking to the two year olds only. No such effect was found in speech to the five year olds (see Figure 2). This interaction effect for subject group and type of listener is shown in Figure 3. The interaction is highly significant ($F(1,22)=108.97$, $p < .001$).

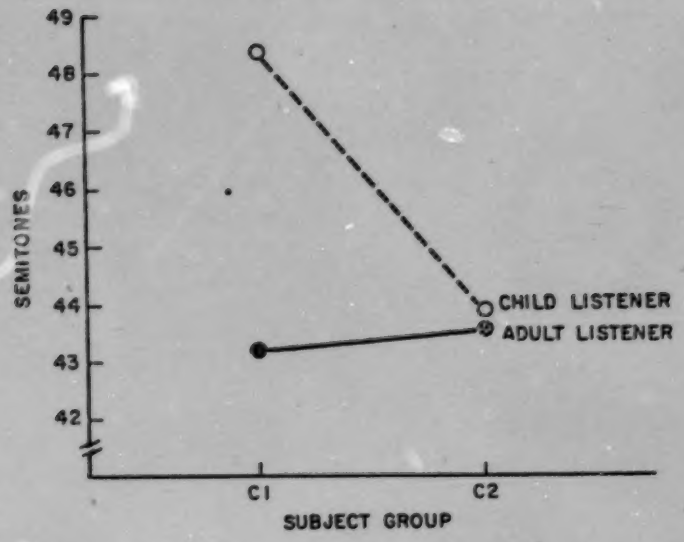


Figure 3. Mean fundamental frequency level by subject group and type of listener.

3.2. Frequency range.

The frequency range for each subject was determined for each session separately. The range was defined by the lowest and highest frequency produced in the session, shown by the figures in Tables 5 and 6. The ranges represent the lowest and highest frequencies actually employed by the subjects in their speech during the session rather than their phonational frequency range. The latter term refers to vocal frequencies ranging from the lowest sustainable tone in the modal register to the highest sustainable tone in falsetto. The ranges for the C1 and C2 subjects in the adult-listener sessions extend from 75 and 80 Hz at the low end to 160 and 185 Hz at the high end. These figures represent a span of approximately 1/2-1 octave (Figures 4 and 5). This range span corresponds well to findings for female speakers by other investigators (Duffy, 1958; Linke, 1953). The mean ranges for the C1 and C2 subjects in the adult listener sessions are quite similar--10.5 and 10.9 semitones. This difference was not significant ($t=.61$; $df=11$; $t_{(.05)}=2.20$).

The frequency ranges of the C1 and C2 subjects in the child-listener sessions were greater than those in the adult-listener sessions. The smallest and largest spans were 200 and 425 Hz among the C1 subjects and 125 and 250 Hz among the C2 subjects. For a number of C1 subjects, the ranges approach a two octave span. The differences between the frequency ranges used to adult-listener versus child-listener for the C1 subjects was highly significant ($t=9.48$; $df=11$; $t_{(.001)}=3.11$). This difference in the range in speech addressed to an adult-listener versus a child-listener was also significant for the C2 subjects ($t=3.376$; $df=22$; $t_{(.001)}=3.11$).

Table 5
Frequency range for C1 subjects by type of listener.

	Adult-listener				Child-listener			
	Lowest/Highest frequency (Hz)	(st)	Range (Hz)	(st)	Lowest/Highest frequency (Hz)	(st)	Range (Hz)	(st)
C101	150/310	38.4/50.9	160	12.5	165/490	40.0/58.9	325	18.9
C102	145/230	37.8/45.7	85	7.9	150/445	38.4/57.2	295	18.8
C103	140/270	37.2/48.5	130	11.3	145/525	37.8/60.1	380	22.3
C104	135/250	36.6/47.3	115	10.7	150/350	38.4/53.0	200	14.6
C105	145/300	37.8/50.4	155	12.6	185/450	42.0/57.4	265	15.4
C106	145/270	37.8/48.5	125	10.7	175/580	41.1/61.9	405	20.8
C107	160/335	39.5/52.3	175	12.8	150/575	38.4/61.6	425	23.2
C108	190/270	42.5/48.5	80	6.0	180/530	41.6/60.3	350	18.7
C109	150/265	38.4/48.2	115	9.8	150/510	38.4/59.6	360	21.2
C110	155/310	39.0/51.0	155	12.0	160/460	39.5/57.8	300	18.3
C111	160/260	39.5/47.9	100	8.4	170/510	40.6/59.6	340	19.0
C112	130/245	36.0/46.9	115	10.9	150/450	38.4/57.4	300	19.0
			$\bar{X}$	10.5			$\bar{X}$	19.2

Table 6
Frequency range for C2 subjects by type of listener.

	Adult-listener				Child-listener			
	Lowest/Highest frequency (Hz)	(st)	Range (Hz)	(st)	Lowest/Highest frequency (Hz)	(st)	Range (Hz)	(st)
C201	130/240	35.9/46.5	110	10.6	150/310	38.4/50.9	160	12.5
C202	140/280	37.2/49.2	140	12.0	140/310	37.2/50.9	170	13.7
C203	145/310	37.8/50.9	165	13.1	150/300	38.4/50.4	150	12.0
C204	145/265	37.8/48.2	120	10.4	140/265	37.2/48.2	125	11.0
C205	160/240	39.5/46.5	80	7.0	140/265	37.2/48.2	125	11.0
C206	155/340	38.9/52.5	185	13.6	150/400	38.4/55.3	250	16.9
C207	150/315	38.4/51.2	165	12.8	140/310	37.2/50.9	179	13.7
C208	145/240	37.8/46.5	95	8.7	150/275	38.4/48.9	125	10.5
C209	155/320	38.9/51.5	155	12.6	150/310	38.4/50.9	160	12.5
C210	150/280	38.4/49.2	130	19.8	150/315	38.4/51.2	165	12.8
C211	180/290	41.5/49.8	110	8.3	150/330	38.4/52.0	180	13.6
C212	140/260	37.2/47.9	120	10.7	130/250	35.9/47.2	120	11.3
			$\bar{X}$ =	10.9			$\bar{X}$ =	12.6

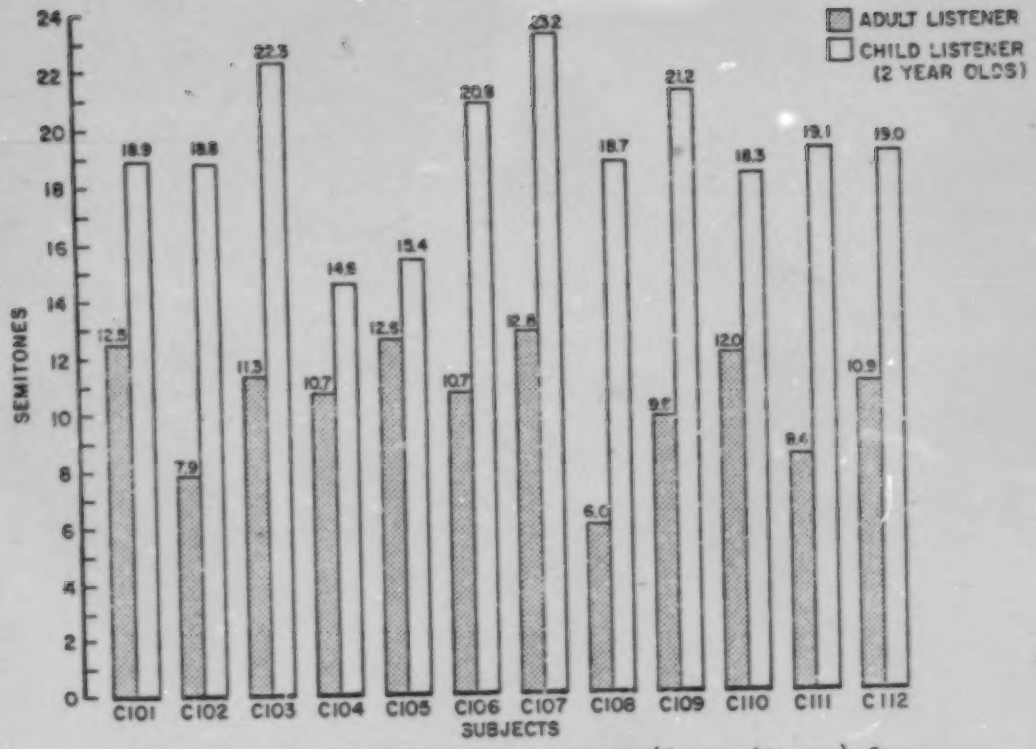


Figure 4. Comparison of frequency ranges (in semitones) for C1 subjects by type of listener.

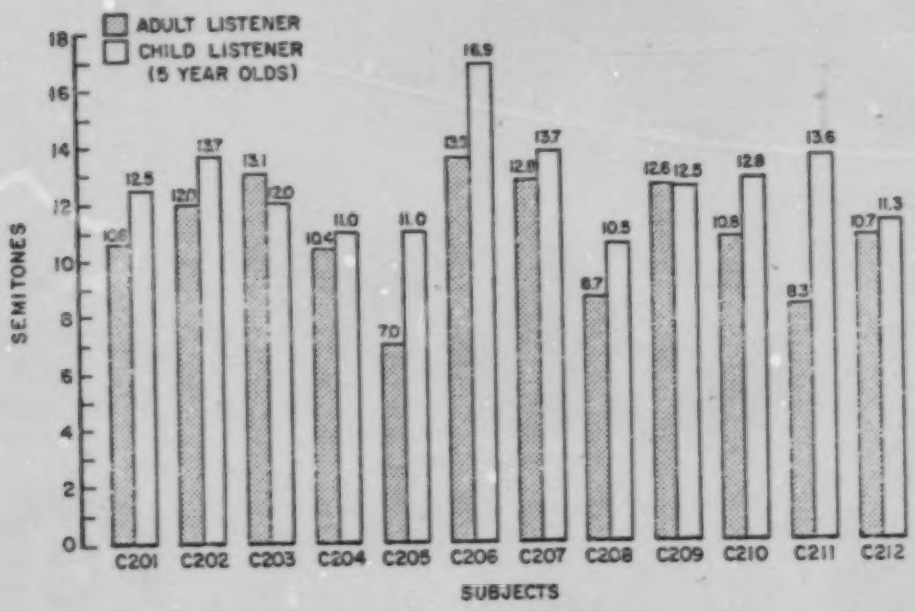


Figure 5. Comparison of frequency ranges (in semitones) for C2 subjects by type of listener.

It is evident from Figure 6 that the low end of the frequency range in speech to the two year old listeners is about the same as it is in the adult-listener sessions. The range is expanded greatly at the higher frequency end. The effect is similar in speech to the five year olds, but the increase in the span is not as large (Figure 7). In addition, the range frequencies for C205 and C211 show another pattern. Here the range in the child-listener session is extended at both the lower and higher end.

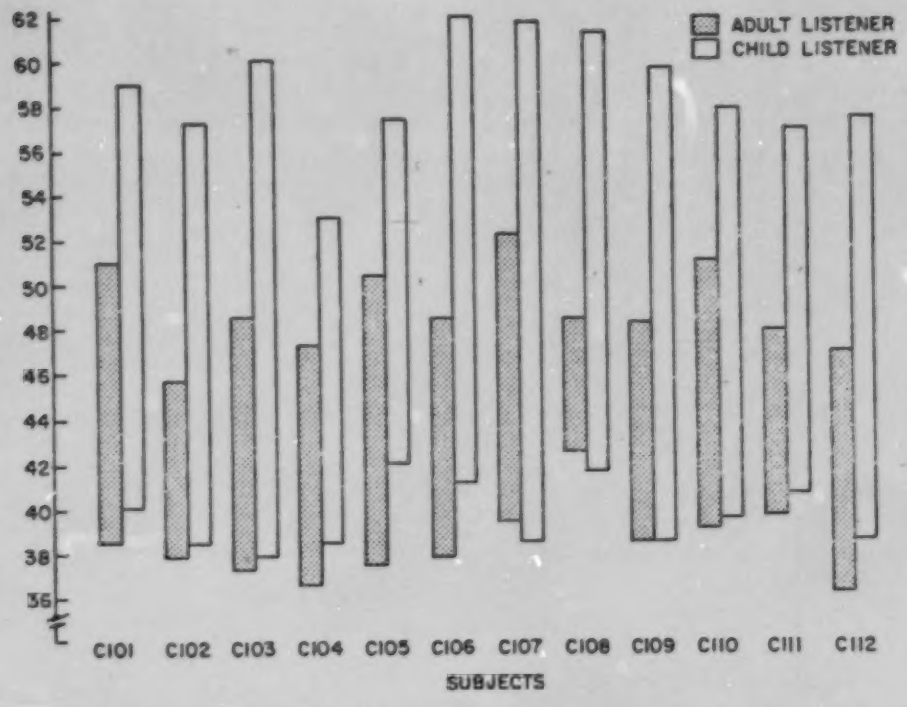


Figure 6. Frequency range for C1 subjects by type of listener.

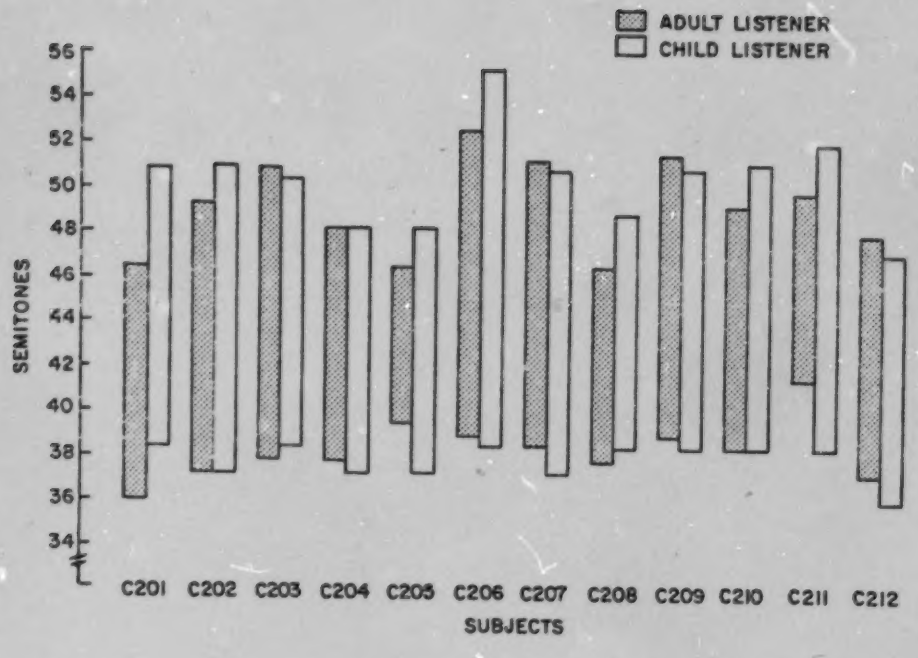


Figure 7. Frequency range for C2 subjects by type of listener.

3.3. Sentence final pitch terminals.

The sentences sampled from the Picture task were declarative and imperative and therefore we would expect these sentences to have a falling final pitch terminal. All the sentences were produced with a falling terminal by the subjects in the adult-listener sessions. However, in the child-listener sessions this was not the case. A prominent feature in speech to the two year olds (C1 subjects, child-listener session) was a rising final pitch terminal in the Puzzle task sentences and occasionally in the Picture task sentences. The number of sentences with a rising final pitch contour for each task is shown in Table 7. Twenty five percent of the sentences spoken by the C1 subjects ended with a rising terminal. All but one subject (C108) used the rising terminal in at least one sentence during these two tasks. This result is highly significant (Fischer sign test, B=11, $b_{.0032}$, 12, 1/2). Furthermore, most of the sentences with rising pitch terminals (85 percent) were produced in the Puzzle task and were therefore all imperative sentences.

Table 7
Number of sentences with rising pitch terminals in the child-listener sessions by subject group and task.

C1 Subjects

	Picture Task (5 sentences)	Puzzle Task (8 sentences)	Total (13 sentences)
C101	1	3	4
C102	0	4	4
C103	0	5	5
C104	0	2	2
C105	3	2	5
C106	1	5	6
C107	1	7	8
C108	0	0	0
C109	0	1	1
C110	0	1	1
C111	0	2	2
C112	0	1	1
	6	33	39

C2 Subjects

	Picture Task (5 sentences)	Puzzle Task (8 sentences)	Total (13 sentences)
C201	0	0	0
C202	0	0	0
C203	0	0	0
C204	0	1	1
C205	1	1	2
C206	0	1	1
C207	0	1	1
C208	0	0	0
C209	1	1	2
C210	0	1	1
C211	0	0	0
C212	0	1	1
	2	7	9

A rising pitch terminal was also present in speech to the five year olds but was much less common. Nine percent of the sentences spoken by the C2 subjects in the child-listener sessions ended with a rising terminal. Seven out of the twelve C2 subjects used the rising terminal in at least one sentence in the two tasks but no subject used it more than twice. This is not statistically significant (Fischer sign test, B=7, b .3872, 12, 1/2). Here again most of the sentences with rising terminals were produced in the Puzzle task. Thus, we find that when the two year old child is the addressee some of the sentences are produced with a rising terminal, even though they are statements and, surprisingly, imperatives. Ordinarily, in adult speech the rising terminal is restricted to questions. The rising terminal was also used occasionally in speech to the five year olds.

3.4. Use of whispering.

An unexpected finding was the preponderance of whispering used by the C1 subjects in the child-listener session. Whispering never

occurred for either C1 or C2 subjects in the adult-listener sessions. Only two C2 subjects (C206 and C208) used whispering in the child-listener session, and then each used it in only one of the sampled sentences. However, nine out of the twelve C1 subjects used whispering in at least one sentence when speaking to the two year olds. The number of sentences (out of 13) in which whispering was used by each subject was: 1 sentence (C106, C112), 2 sentences (C105, C110, C111), 3 sentences (C101, C102, C109) and 4 sentences (C107). In only one case was an entire sentence whispered. Most often the last half of the sentence contained the whispered syllables. A check of the complete transcripts of the child-listener session for the C1 subjects revealed that the use of whispering was not restricted to the sub-sample chosen for acoustic and perceptual analysis but was evident throughout the entire session; in some cases, the use of whispering was more extensive in the unanalyzed portions of speech. In the case of the three C1 subjects that did not use whispering in the sentences analyzed, all of them made at least some use of whispering at other points in the child-listener session. Of the C2 subjects, only the two subjects mentioned above made any use of whispering in the child-listener session. The use of whispering by and large seems to be restricted to speech to two year olds.

3.5. Duration

Two content words from each of six sentences in the Puzzle task were chosen for the comparison of average duration of syllables between the adult-listener and child-listener sessions. The six sentences (1, 3, 5, 6, 7, 8--see 2.2.2) contain both a verb (push in/take out) and a color term (green/red/blue/orange/purple/brown). The average duration of the syllable nucleus (in msecs) was computed for each subject by session. Verb and color terms were tabulated separately for Tables 8 and 9. In computing the average durations for the verbs, the two word sequence (e.g. push in) was considered as one and the durations were added together. In the case of purple, the only color term pronounced with two syllables, the two syllables were also considered as one item. The color term orange was always pronounced as one syllable, e.g. [arndʒ].

Table 8
Average duration (msec) of verbs in Puzzle Task

	C1 Ss			C2 Ss	
	Adult listener	Child listener		Adult listener	Child listener
C101	263.33	301.67	C201	260.00	236.00
C102	190.00	243.33	C202	286.67	275.00
C103	193.33	183.33	C203	205.00	228.00
C104	225.00	266.67	C204	266.67	266.33
C105	265.00	296.67	C205	190.00	201.67
C106	205.00	238.33	C206	195.00	186.67
C107	176.67	203.33	C207	218.33	211.67
C108	191.67	318.33	C208	198.33	198.33
C109	148.33	181.67	C209	166.67	167.67
C110	210.00	208.33	C210	223.33	236.67
C111	193.33	230.00	C211	198.33	261.67
C112	270.00	315.00	C212	268.33	275.00
	$\bar{X}$ = 210.97	$\bar{X}$ = 248.89		$\bar{X}$ = 223.06	$\bar{X}$ = 228.72

Table 9
Average duration (msec) of color terms in Puzzle Task

	C1 Ss			C2 Ss	
	Adult listener	Child listener		Adult listener	Child listener
C101	238.33	310.00	C201	246.67	251.67
C102	210.00	243.33	C202	210.00	243.33
C103	183.33	223.33	C203	226.67	321.67
C104	190.00	192.00	C204	255.67	296.67
C105	285.00	355.00	C205	181.67	205.00
C106	200.00	218.33	C206	208.33	260.00
C107	205.00	271.67	C207	201.67	230.00
C108	205.00	301.67	C208	205.00	240.00
C109	188.33	218.33	C209	141.67	187.50
C110	205.00	238.33	C210	251.67	271.67
C111	205.00	361.67	C211	200.00	296.67
C112	255.00	398.33	C212	273.33	355.00
	$\bar{X}$ = 214.16	$\bar{X}$ = 277.67		$\bar{X}$ = 216.86	$\bar{X}$ = 263.26

The results for the verbs and color terms will be discussed separately. For the verbs, the difference between the means for the C1 and C2 subjects in the adult-listener session was fairly small (210.97 and 223.06 respectively). This difference was not significant

(Wilcoxon rank sum test, $W^* = 0.87$, $z_{(.05)} = 1.645$). However, for ten out of twelve C1 subjects the average duration of the verbs was higher in the child-listener session than in the adult-listener session. Only six of the C2 subjects showed a similar difference. The difference between the durations in the adult-listener and child-listener session for the C1 subjects was significant (Wilcoxon signed rank sum test, $I^* = 2.30$, $z_{(.05)} = 1.645$) and this is reflected in the difference between the means for the two sessions--210.97 and 248.89. The difference between the means for the two sessions for C2 subjects was smaller (22.306 and 228.72) and was not significant [Wilcoxon signed rank test, B = 6, (.6128, 12, 1/20)).

The results for the color terms are slightly different. The difference between the means for the C1 and C2 groups in the adult-listener session (214.16 and 216.86) was very small and not statistically significant (Wilcoxon rank sum test, $W^* = 0.52$, $z_{(.05)} = 1.645$). However, for all the subjects there was an increase in average duration of color terms in the child-listener session. The difference between the means for the two sessions for both C1 and C2 subjects reflects this fact (C1 subjects--214.16 versus 227.67;C2 subjects--217.86 versus 263.26). Both are significant (Wilcoxon summed rank test, $I^* = 3.45$, $z_{(.001)} = 3.09$).

One factor which has not yet been considered in this analysis is the perception of differences in duration. It is not the case that a one millisecond difference between two stimuli will be perceived and this should be considered in interpreting the results. In fact, the just noticeable difference (jnd) for duration increases as the duration of the standard stimuli increases. There are several studies on the perception of durations but they do not seem to agree on the size of the jnd for different duration values (Stott, 1935; Henry, 1948; Ruhm et al. 1966). However, if we approximate conservatively from the available information (i.e. that 30 msecs is the jnd for durations of 150-200 msecs, 35-40 msecs for durations 200-250 msecs, and 45 msec for durations for 250-300 msec), we should have enough information to correct for the potential effect of the perceptual factor in evaluating differences in duration between the adult-listener and child-listener sessions. For C1 subjects we find that the difference between average duration values for the adult-listener and child-listener sessions are larger than the perceptual threshold. This is the case in all instances for the verbs. For the color terms, this is the case for all subjects except C104 and C106. Therefore, taking the perception factor into consideration does not change the results for the C1 subjects. The original results also hold for the C2 subjects. For verbs, the difference between the adult-listener and child-listener sessions is not significant. Taking the perceptual factor into consideration reduces the number of changes from adult-listener to child-listener to one subject (C212). For color terms, the perceptual factor likewise does not alter the results since only two subjects (C201 and C209) are affected. Thus, we find that the duration of verbs and color terms is significantly longer in sentences spoken to the two year olds. Only

the duration of the color terms is longer in sentences spoken to the five year old children.

3.6. Distribution of primary stress

One result of the perceptual analysis was the finding that primary stress placement was different in speech directed to the child-listener than to the adult.[1] This occurred only in speech directed to the two year olds. The difference observed was the appearance of two primary stressed syllables in a sentence which ordinarily, in adult-adult communication, would contain only one primary stress. The sentences in which this phenomenon was noticed were the six sentences of the form "Push in _____" and "Take out _____" in the Puzzle Task.

As shown in Table 10, there were only scattered instances of use of double primary stress in speech directed to the adult listener. Five out of 144 sentence samples fall into this category, amounting to approximately three percent of the sentences directed to the adult listener in the Puzzle task. Only three of the C1 subjects (C101, C105, C112) and two of the C2 subjects (C202, C207) showed use of double primary stress. This stress distribution appeared in only one sentence for each of these subjects.

Table 10
Number of instances of two primary stresses per sentence in six sentences from the Puzzle task by type of listener.

	C1 Subjects			C2 Subjects	
	Adult listener	Child listener		Adult listener	Child listener
C101	1	2	C201	-	-
C102	-	1	C202	1	2
C103	-	-	C203	-	-
C104	-	3	C204	-	-
C105	1	3	C205	-	-
C106	-	2	C206	-	-
C107	-	1	C207	1	-
C108	-	1	C208	-	-
C109	-	-	C209	-	-
C110	-	1	C210	-	-
C111	-	1	C211	-	-
C112	-	3	C212	-	-
	3/12 Ss	10/11 Ss		2/12 Ss	1/12 Ss

The situation is somewhat the same in speech directed to the five year olds. Only one C2 subject (C202) used double primary stress in the child-listener session. On the other hand, ten of the C1 subjects used double primary stress in at least one sentence (Fischer Sign test, B = 10, b.0193, 12, 1/2). Of these, three subjects used double primary stress in each of three sentences from the Puzzle task, two subjects in

two sentences each, and five subjects in one sentence each. This represents a total of 18 out of 72 sentences or about 25 percent of the Puzzle task sentences directed to the two year old child listener. Only two C1 subjects (C103, C109) did not exhibit this characteristic in their speech.

It seems evident, therefore, that the assignment of more than one primary stress to the short and simple sentences of the Puzzle task occurs almost exclusively in the case where the sentences are directed to the two year old. Only scattered instances occur in speech directed to the adult listener and the five year old.

3.7. Summary of results.

Six major analyses were performed on samples of speech directed to adult listeners and to child listeners. Speech to the two year olds differed on the six analyses from speech to the adult listeners. Only some of these differences were found between speech directed to the five year old child listener and the adult listener.

The results indicate that:

(a) The average fundamental pitch of the speaker's voice is higher to the two year old than to the adult. This was not the case for speech to the five year old.

(b) The frequency range of the speaker's voice is greater to the two year old and to the five year old in comparison with the speech range to the adult listener. The expansion occurs at the high end of the range.

(c) Speech to the two year old contains many instances of rising sentence final pitch terminals in sentences where the grammatical form would normally dictate a falling pitch, e.t. imperatives. This feature is absent from speech directed to the adult listener and to the five year olds.

(d) Whispered or partially whispered words appear in speech directed to the two year olds. This characteristic is absent from speech directed to the adult listener and to the five year old.

(e) The duration of certain content words in selected sentences is prolonged in speech to the child listener as compared to that of the adult listener. In the Puzzle task sentences, the verbs and the color terms had longer durations in speech to the two year olds than to the adults. Only the durations of the color terms were so affected in speech to the five year old.

(f) Speech directed to the two year old contains many cases of more than one instance of primary stress assigned within a sentence unit. This feature is absent in speech directed to the five year old and to the adult listener.

Footnote to Chapter 3.

1. Two transcribers independently transcribed primary stress placement in these sentences. A correlation of +.92 was found between the transcribers.

CHAPTER 4: DISCUSSION

4.1. Characteristics of the baby-talk register: functions

It has been shown in this and other studies that speech directed to young children differs in systematic ways from speech directed to an adult. Given that these findings are reliable and that these differences do exist, the question that arises is what functions these particular features serve. Of what potential use is this "special" speech to the child learning language? I would suggest that the various features of the baby talk register can serve at least two functions--an analytic function and a social function.

A child learning language is constantly analyzing the speech that he hears as a means for learning the rules of his language. He then uses these rules to effectively communicate with those around him. This analytic endeavor is no doubt an enormous task. Although some theorists, notably Chomsky (1965), have suggested that language input of almost any kind is sufficient for language learning, this seems unlikely. Speech directed to children is rich in redundancy, repetition and other devices which multiply the linguistic information available in the utterance. A plausible hypothesis is that characteristics found to be prevalent in the baby talk register have a distinct function in the child's analytic endeavor. They assist the child in delimiting sentences, words and other syntactic constituents and possibly do more. Among the characteristics of this register that have such an analytic function are some of the aspects of prosody discussed in the previous chapter. This is because the prosodic features of speech are the primary means by which a speaker organizes units above the level of the phonological segment into groups.

Knowing specific linguistic rules, however, such as those governing proper word order is not enough. To communicate effectively the child must know, among other things, the rules for how to engage in a verbal exchange with another person. And, in order to interact with a child, it is necessary to gain the child's attention for conversation. Getting the initial attention of your interlocutor in a conversation is a primary prerequisite to beginning a communicative exchange. Keeping the attention of your interlocutor is necessary for the maintenance of communication. It is hypothesized that some features of the baby-talk register serve a social function, i.e. to initiate and maintain communication between adult and child (and likewise between an older child and a young child).

Prosodic features can cue the child to pay attention and listen to the speech of the person attempting to communicate with him. Other verbal means used include frequent repetitions of the child's name. The features which serve a social function complement those features which assist the child's analytic endeavor, since a child must attend to a particular set of speech in order to utilize whatever analytic cues are provided.

In the following two sections the findings of the present study will be considered in terms of the above two functions. It will

become apparent that some features may play a dual role, that is, simultaneously serve a social and an analytic function. Other features seem to serve predominately one function or the other.

4.1.1. Prosodic features: social function

The higher pitched voice used by the subjects in this study to the two-year-olds can be viewed as serving primarily a social function. The higher pitch is quite unique to this function. It may in fact be the most salient characteristic that serves to mark and thus set apart the baby talk register from other registers. Even the most casual of observers seems to notice it. An utterance spoken with a higher pitched voice marks that message as intended for the child listener. The message may in other respects be 'tailored' to the language abilities of the child. A message so marked prosodically is foregrounded against the background of adult-adult communication.

The question arises of whether the higher pitch level is in some way more salient to the child. There seems to be little if any evidence that the young child's hearing apparatus is more sensitive to higher frequencies than lower ones. It is the case, however, that in general higher frequencies are more audible than lower frequencies, given the same intensity level. The higher pitch level is, of course, closer to the pitch of the child's own voice. One can only speculate as to whether this factor plays a role. In any event, it is at least plausible that a higher pitch level serves a social function by regulating communication with the child. It attracts the child's ttention to verbal material directed to him.

The expanded pitch range observed in the speech of the adult subjects to both groups of child listeners (two year olds and five year olds) also has a similar function. The extension of the range was primarily in the upward direction. The presence of high pitch peaks in utterances intended for the child listener, may be salient cues that mark particular sections of a speaker's speech and therefore make them stand out. The finding of higher pitch in speech to the child listener should be considered with the understanding that such speech is not characterized by high levels of pitch in every syllable nucleus. It is rather the case that the peaks in a sentence unit that do appear are in many cases exaggerated in comparison with speech to adults.

The use of whispering also may be considered as an example of the social function of prosodic characteristics. Whispering, in fact, is very closely allied with the extension of the range capabilities of speech. Whereas the range is expanded at the higher end by the presence of higher syllable peaks, an extension at the lower end of the range may result in the voice going into whisper. There is ample evidence from languages using tone that when a speaker produces an exaggerated rendition of a low tone, a whisper may result. In Serbo-Croatian, for example, Ivić and Lehiste (1969) observed that the voices of speakers who exaggerated the low-to-high tone at the end of an utterance went into whisper on the low tone portion of the

utterance. In some African languages when there is a lowering of tone at the end of questions, whisper often appears (Will Leben, personal communication). Thus it seems that the expansion of range in the baby talk register occurs at both the high and low end of the voice range. Both the high pitched syllable and the whispered syllable stand out and perhaps have attention getting properties.

Finally, the preponderance of rising sentence final terminals in speech to the child listener may serve a social function--to regulate conversation between adult and child. The predominance of rising terminals may cue the child as to when he is expected to respond, since the question is the grammatical form most often associated with a rising terminal, and questions normally demand an answer. Also, it has been noted that sustained or rising pitch in place of terminal falls is generally used to indicate "unfinished business." (Bolinger, 1961).

It is not uncommon to observe an adult asking a child listener a question and then answering the question if the child does not respond to complete the exchange. The completeness of the question/answer sequence in terms of a communication unit is best seen if one thinks of the question forming the first half of a contour (ending with a rising pitch terminal) and the answer continuing the contour and ending with a falling pitch terminal signals completion of the contour and simultaneously the completion of the exchange. The presence of many rising pitch terminals may serve, then, not only to regulate the conversation between adult and child but also to keep the child's attention. One must pay attention in a conversation in order to know when it is one's turn to speak.

Thus, the higher pitch level, the expanded pitch range, the use of whispering, and the predominance of rising pitch terminals in adult speech directed to a child listener can be interpreted as serving a social function.

4.1.2. Prosodic features: Analytic function

Some of the prosodic characteristics discussed in the previous chapter seem to serve an analytic function. One feature which plays a dual role is the preponderance of rising pitch terminals. These may be used to cue the child to the location of sentence boundaries. The fact that a high pitch is attained at the end of the sentence (the boundary) is significant because the high pitch would tend to accentuate the termination of the sentence by the speaker. Furthermore, the rising pitch terminals were associated with sentences which by grammatical form were imperatives and not interrogatives. It is unlikely that the adults were using the imperatives as questions since the context of the sentences was the administration of instructions to a task. Making the imperatives into questions would indicate that the speaker was unsure of the instructions. This never occurred even in the adult-adult sessions when the subject was completely unfamiliar with the task and the instructions she was to administer. It is more likely that the rising pitch terminals on sentences of the imperative

form functioned as a signal both to regulate the verbal exchange (social function) and to mark the sentence boundary (analytic function).

The longer durations of certain words (one or both content words in the sentences studied) can also be seen as potentially having an analytic function. Duration is an important correlate of stress, although there is no direct, one-to-one relationship between the duration of a syllable and the degree of stress it carries. For the following discussion, the situation will be somewhat simplified by disregarding other factors involved.

In speech to the five year olds the durations of the color terms were significantly greater than those directed to the adult listener. The extension of duration in color terms can be viewed as a way to supplement the function of contrastive stress on the unit. By prolonging the duration of the syllable nucleus of red in Push in the red piece the speaker implies with greater force the propositions "not the yellow piece, not the blue piece."

In speech to the two year olds, the duration of both the color terms and the verbs were greater than to the adult listener, indicating both emphatic stress on the verbs and contrastive stress on the color terms. This may be the "key" words in the sentence. These were the only words the child needed to understand in order to carry out the command correctly. For example, in the case of the first sentence Push in the green square, the listener had only to attend and understand the words push and green to correctly complete the demanded action. The word square is redundant here since there were no other green pieces.

The longer durations of verbs and color terms to the two year olds no doubt contributed greatly to the perception of two primary stresses in many sentences directed to them. When two primary stresses were transcribed, they were marked as falling on the verb and color term of the sentence. Aside from the above mentioned function (to indicate keywords), two primary stresses may serve to divide up a sentence perceptually into smaller units. The adult thereby segments the sentence into pieces he/she thinks are of adequate size for the child to process easily. The same sentence which, when directed to the adult, would normally contain only one primary stress, would be divided into multiple units for the child. Furthermore, it is interesting that the chunks that sentences such as Push in the green square are divided into are the major constituents of the sentence. By this division, the adult may be providing the child with important information about constituent structure. A look at more sentences with different structures would be necessary before anything more than a tentative statement on this point could be made. However, my initial inspection of all the sentences contained in the Rainbow Passage, as read to the child listener, confirms this position. In reading to the child listener, (to the two year olds in particular) the longer sentences are divided up prosodically (here with respect only to stress) into smaller units. These smaller units are in most cases the major constituents of the sentence.

Finally, the longer durations of the color terms and verbs, and the extra primary stress may function to teach the child how to systematically mark emphasis in his own speech. This is undoubtedly secondary to the direct communicative benefit of these features, but it is something that must be learned at some point since languages differ in the ways in which they express emphasis.

4.2. Age of child listener

As evidenced by the various prosodic differences found between speech to the two year olds and to the five year olds, some of the devices which are commonly used to the first group have disappeared or are greatly diminished in the speech to the second group. A summary of all the characteristics studied and their presence as a function of age appear in Table 11. These observed changes are no doubt due to the linguistic maturation of the child. The age groupings are only meant to be rough indicators of the linguistic abilities of the child and should be considered as such.

The prosodic features that were seen to serve primarily a social function disappeared earlier from the speech of the adults than those serving an analytic one. These include higher fundamental pitch, the use of whispering, and the use of rising final terminals in sentences of imperative form. This is explained by the fact that by the time most children reach the age of four or five, their attention span has improved greatly, eliminating the need for attention getting and attention holding devices on the part of the speaker. Furthermore, by this age probably all of the children have learned the rudimentary rules of conversational exchange and some have already become masters of more sophisticated conversational skills such as verbal manipulation.

Table 11
Presence and absence of some prosodic characteristics in adult speech by function and age of child listener.

Characteristic	Function(s)	Age of Child Listener	
		Two year olds	Five year olds
Higher fundamental pitch	Primarily social	YES	NO
Expanded frequency range	Primarily social	YES	YES
Use of whispering	Primarily social	YES	NO
Rising sentence final pitch terminals	Social/Analytic	YES	NO
Longer durations of verbs	Primarily analytic	YES	NO
of color terms		YES	YES
Use of two primary stresses per sentence unit	Primarily analytic	YES	NO

One feature which does seem to remain is an expanded frequency range, indicating that some instances of high pitch do appear in utterances directed to the five year olds. Not all of the mothers' speech directed to the five year olds exhibited this characteristic. Those that did are in the minority; and these subjects' speech exhibited high pitch peaks in only some utterances. The reasons for these differences between individuals is difficult to determine. My observations indicate that the occasional use of higher pitch to the five year old has little to do with the child's verbal abilities. Instead, it seems to be determined primarily on how interested the child is in the task at hand. When the five year old child is distracted from the task, some mothers will use higher pitch as a device to bring the child back to the task.

By the time the child is five years old, many of the features hypothesized to contribute to the child's analytic endeavor have either disappeared or are greatly reduced in frequency. The emphatic stress on the verbs in the Puzzle task sentences and the use of two primary stresses per sentence are among these. Since the five year old is producing, and therefore presumably fully comprehends sentences of the type in the Puzzle task it is no longer even potentially useful to "cue" the child to the "key" words (the verb) as was necessary earlier. Also, the utterance need not be divided up into such small units as before. The speakers, however, still feel it necessary to modify their speech by prolonging the duration of the color terms in the Puzzle task.

4.3. Speaker interviews

Interviews were conducted with the subjects who participated in the study. One purpose of the interview was to determine if the speakers were aware of the modifications they (and other adults) make in speech to the child listener, and the kinds of differences they noticed. A second purpose was to determine what kinds of verbal means the speakers thought that they used if they specifically wanted to gain the child's attention. It was of interest to see how closely the respondents' answers to the latter question would match their answers to the former and how these related to the findings of the study.

The respondents were all very cooperative and in most cases volunteered a great deal of information about their observations. Many of their answers were well thought out and quite perceptive. One unfortunate drawback in this type of interview, however, was that the respondents did not possess the vocabulary necessary to make precise statements about such matters. As a result, they used terminology which was sometimes vague and ambiguous. For example, many mothers noted that they felt their speech to the child listener was more animated than to the adult, and that their voice was filled with excitement. It is difficult to determine the precise parameters involved when such labels are used (for an excellent discussion of the problem see Crystal 1969). Despite such drawbacks, some statements

made by the respondents are of potential value not only to the interpretation of the experimental results of the study but to the clarification of certain broader issues that are at the moment vague. Data bearing on such matters have been extracted from the interviews and are discussed below.

Most of the subjects noticed differences in their speech as a function of the type of listener (child versus adult) but not all to the same degree. The subjects who addressed the two year olds said they definitely noticed changes in their speech and observed that the differences were quite dramatic. On the other hand, most of the subjects who addressed the five year olds were more tentative about their responses and qualified them with such words as "probably", "perhaps", and "sometimes". Three of these subjects said they doubted that there were any differences at all, but added that when their child was younger they had noticed differences in their speech. The observations of these two groups of subjects are consistent with the findings in this study, that relative to speech directed to the adult, there are many more differences in speech to the two year old than in speech to the five year old.

The kinds of changes that the C1 subjects noted were higher pitch of voice (C102, C103, C105, C107, C110), an expanded range (C103, C105), less volume in voice (C104, C105, C107, C112) and slower speech (C112). Some of these factors were investigated in the present study and found to be empirically valid. Three of the changes noted by the C1 subjects--higher pitch (C203), decreased volume (C202, C212), and slower speech (C206, C206, C208, C210). All the C2 subjects, however, qualified their observations by specifying the particular situation which would trigger these modifications in their speech. One subject (C203), for example, reported that she used a higher pitched voice to her five year old, but only when she was upset about something the child had done and was scolding him. All the C2 subjects who observed differences emphasized that the changes in speech to the five year old were not present at all times. Rather, the changes occurred in particular situations. The most common situations mentioned were: (a) when the child was in a certain state--tired, emotionally upset; (b) when the adult had to restate a request or command, etc., after failure to convey the message adequately; and (c) when the adult was presenting new information to the child, which might be difficult to understand or carry out.

The C1 subjects, on the other hand, admitted using the features they noticed almost always when they were interacting verbally with the two year olds. They reported noticing these characteristics most when they were trying to get the child's attention or when they were in a one-to-one interaction with the child. They believed that they used the features in their speech because it got them results, i.e. it helped make their communication with the child more effective. Several subjects noted that when they failed to use these features, their communication was not understood by the child. As one subject put it,

> There are plenty of times I don't stop to think that he's two and I'll just mumble something at him or make some kind of demand on him and don't really think about whether or not he can understand it. And that's when he's most likely not to respond at all. (C106)

All the subjects felt that using the features (higher pitch, etc.) got them results in communicating with the child and admitted using them in their speech, although one half of the speakers said it was probably not good for the child's language development to "talk down to them" in this fashion.

The most commonly mentioned factor triggering the use of these features of the baby talk register was cognizance of the child's presence. This is not a novel observation (Snow, 1972). Certain subjects also mentioned the specific factors that for them determined whether they used the registral features or not in a particular instance. Age and physical size determined the initial level at which an adult would begin an exchange with a child. Several subjects reported occasions on which they spoke to a child of small stature using what they thought to be appropriate features, only to find their speech inappropriate to the child's actual linguistic capability. They claimed making an immediate adjustment in their speech to the verbal abilities of the child. Feedback from the child plays a crucial role in determining the particular registral features selected by the speaker.

Although all the subjects interviewed had something slightly different to say about the devices they used to get a child's attention, they all agreed that changing one's voice so that it would maximally contrast with the ongoing level of speech was the most effective means.

> Often I find I have to do something clever to get a young child's attention. And its more effective to do something competely ridiculous or out of the ordinary. Anything that departs from the ordinary or expected. That gets their attention best. (C202)

The devices mentioned ranged from raising the pitch of the voice, talking louder, and wide variation in pitch on one end of speaking softer, slower, and using whispering. The latter approach was associated with a particular style of dealing with the situation.

> It's the background I've had. My mother did that to me too. If she really wanted to get my attention she always whispered. I remember that my sister and I always sat up and took notice. I find it works with my children also. (C205)

These persons were in the minority. Most mothers chose the other approach.

It seems then, in general, that foreground[illegible]he speech directed to the child using the devices that produce con[illegible] was what the mothers found through experience to be most eff[illegible]ve with young children. This would uphold the interpretation suggested above that some of the prosodic features serve a social function, i.e. to attract and hold the child's attention.

4.4. Implications for further research.

The present study is a preliminary investigation of the prosodic and paralinguistic characteristics of speech directed to children. It only begins to explore the parameters involved. This section presents some opinions on the important issues that need to be explored as subsequent steps in this ongoing investigation. The discussion which follows will be in two parts. The first part suggests further analyses of potential value on the specific material which formed that data for the present study. In it I suggest other variables that could and should be investigated. The second part offers some questions that must be answered as a next step in exploring this particular line of research. Major areas of future research are indicated and a schematic discussion is presented for gathering experimental evidence.

Further analyses. Several factors are not considered in the present study because it is an initial effort in a virtually unexplored area. For one, further investigation of pitch variability is necessary. Frequency range is a limited indicator of variability because it provides no information on the distribution of pitch and is greatly affected by a few extreme cases. Frequency distribution curves for each set of speech samples would provide some evidence on this factor.

Another factor which warrants further study is speech rate. The subjects indicated in the interviews that they spoke slower to the child listener than to the adult. Two measures might suffice here: average number of words per minute for a selection, e.g. the Rainbow passage, and rate of speech during continuous flow of speech.

Finally, more perceptual analyses of the data need to be performed. Specifically, the utterances directed to the child need to be transcribed so that a study could be made of the types of intonation patterns that are used in speech to the child listener, the frequency with which certain patterns are utilized, and the unique features (if any) that are employed. A number of different transcription systems are available. Some of them would undoubtedly be satisfactory for this purpose. There is a fund of information that could be gained from such an analysis, which would also be very useful to the student of the acquisition of non-segmental phonology (Crystal, 1969a).

Further issues. There are at least two questions that, in my opinion, merit further study. The first concerns the identifiability of the baby-talk register. Are adults able to identify a stretch of speech as directed to a child listener or to an adult listener? And, specifically, are the prosodic and paralinguistic features of the

baby-talk register sufficiently salient cues to form the basis for such identification? A study has been begun to answer these questions. It consists of two parts. In the first part subjects listen to short excerpts of speech. Some of these excerpts were taken from speech sequences spoken to an adult listener, and others from sequences spoken to the child. The excerpts are free from all extraneous cues (e.g. the child's voice in background, etc.) except those which are characteristic of the register involved, (e.g. use of special lexical items, repetitions, shorter sentences, etc.). The subjects are asked to label each speech excerpt as "adult listener" or "child listener." In the second part, a different set of subjects listen to repeated presentations of sentences from a list which they have read previously. The sentences are produced by several different speakers. Some of the sentences were recorded when the speaker was addressing the sentence to a child, others when the speaker was addressing an adult. Each pair of sentences is selected so that the segmental aspects of the sentence match one another very closely and that only the prosodic and paralinguistic aspects differ. The sentences are presented to the subjects in random order. The subjects' task is to label each sentence as to whether it was spoken to a child listener or to an adult. The purpose of the latter part of the study is to determine if subjects can identify the age status of the listener to whom speech was directed on the basis of prosodic cues alone.

It is not enough, however, to show that there are prosodic differences between speech directed to a child listener and an adult, and that, furthermore, these differences are salient to the degree that adults can recognize whether excerpts of speech are child directed or adult directed on the basis of these cues. The second question, therefore, concerns the effect, if any, of such prosodic features on the child's attention and performance. One possible procedure for investigating this question would be to place a loud speaker into a large toy animal and have this animal "speak" to the child. The animal could, for instance, give the child directions on completing a task, e.g. putting together a picture puzzle. In one condition, the sentences given as directions would be sentences prerecorded in an adult listener situation. In another condition, the sentences would be those pre-recorded in a child listener situation. The child's response, or the lack of it, would be compared as a function of condition. It is hypothesized that the child would respond more often and more appropriately to directions produced with baby talk characteristics than to the other sentences.

CHAPTER 5: CONCLUSION

5.1. Conclusion

As stated in the introduction, the goal of the present investigation is not only to provide some empirical data on a particular aspect of the linguistic characteristics of speech addressed to young children but also to consider the broader implications of such findings for the theory of language development and linguistic theory more generally. In what follows, the findings of the study will be briefly summarized and discussed in the light of these broader considerations.

5.2. Implications for theory of language development

The first question the present investigation was designed to answer is primarily descriptive: What are the linguistic characteristics of speech to young children? Specifically, what are the prosodic and paralinguistic features which most commonly appear in speech to the young child? Further, how does the use of these features change as the child gets older and thus more sophisticated linguistically? The analysis of speech samples of mothers addressing their two year old or five year old child indicated that speech to the younger children contained (1) higher mean fundamental frequency, (2) a greatly expanded frequency range, (3) numerous instances of rising sentence pitch terminals in declarative and imperative sentences, (4) a high incidence of whispering, (5) increased durations of certain key content words in sentences, and (6) the multiple assignment of primary stress within a sentence unit. Speech to the older children, however, contained only features (2) and (5) and even then the frequency range was less expanded with only one group of words showing longer duration. Although this list of features is preliminary and incomplete it is clear that the adult speaker makes a systematic adjustment in the prosodic and paralinguistic aspects of speech addressed to the young child and that these adjustments vary in relation to the age of the child listener.

With respect to a theory of language development a relevant question which one would pose concerning the characteristics of speech directed to children is whether any of these features of the verbal environment are in some way critical for language learning. Since a definitive experiment to answer this question is not possible on ethical and moral grounds, one can only speculate with greater or lesser assurance, as to whether one feature or another (or more likely some combination of features) may play a significant role. Such speculation would best be directed toward analysis aimed at exploring the function(s) certain features may serve in speech directed to children.

In this investigation the features found in speech to the child listener were classified according to whether they could be interpreted as serving primarily (a) an analytic function, to aid the child in the linguistic analysis of the speech he hears, or (b) a social

function, to capture and maintain the child's attention to the speech directed to him. Features serving a primarily social function were higher fundamental frequency, expanded frequency range and use of whispering. On the other hand, longer durations of content words and assignment of multiple primary stress within a sentence unit were interpreted as serving primarily an analytic function. The characteristic use of rising sentence final pitch terminals was seen to serve both functions.

There is, in my opinion, a stronger argument for considering features serving an analytic function as playing a "critical" role in language acquisition. The social function, after all, can be adequately fulfilled by other means of communication. For example, the use of kinesic cues, i.e., touching, using facial expressions, etc., may be just as effective in capturing and maintaining the child's attention. The features serving primarily an analytic function may not only provide the child with cues necessary for the child's analysis of constituent structure (and the like) but may also be invaluable to the child's learning of certain aspects of the prosodic system of the language he is learning, such as, for example, the system of contrastive stress in English. However, this is not directly true for intonation contour per se. The child has to learn that different utterances (e.g., imperative, declarative, questions) do contrast in adult speech.

Finally, the search for "critical" features for language acquisition may be a misguided venture in itself. What is strongly implied in the question of whether one or another feature is critical is an all-or-none situation. I would suggest that it is more likely that some combination of features is necessary for language acquisition and that each individual feature is only more or less suitable or effective in achieving the goal. Thus, the choice of one subset of effective features may produce a desired result just as effectively as another. As evidenced in the range of individual differences among the adult speakers in the present study, the same effect, e.g., getting the attention of the child, was accomplished in a variety of ways.

5.3. Implications for linguistic theory.

It has been suggested that the study of "simplified registers", of which the baby talk register is one, can contribute to a general linguistic theory by clarifying the notion of simplification in language and the elucidating factors which govern language use (Ferguson 1972). We can now examine how some of the features investigated in the present study could contribute to these endeavors.

Assuming that the hypothesis of universal simplification processes is a productive one, such features of the baby talk register as the prolongation of "key" content words in sentences can be viewed as leading ultimately to the simplification of structure. Longer duration is generally highly correlated with stress placement. Elements that ordinarily carry reduced stress are more often eliminated in the simplification process. For example, articles are often omitted. Thus a sentence type used in the present investigation, e.g., "Push in the blue piece," was consistently produced in such a way that the article "the" was almost totally obscured, while the meaning of the

command was kept intact. Evidence such as this can perhaps be used to explain the omission of the article in languages thought to have undergone some simplification processes, for example, pidgins.

A comparison of the use of prosodic features in the different registers could provide a better understanding of which factors govern the use of particular features. For example, the use of a high pitched voice is a feature peculiar to the baby talk register. None of the other "simplified registers" show a use of this feature. Whatever the reasons for this fact, it seems evident that the presence of a child addressee seems to play a role in the use of this feature. These are only some examples, and perhaps superficial ones at that. As information on "simplified registers" and other registers increases, the contributions of such studies to linguistic theory will become more evident and the contribution to its development should be more widespread.

References

Avram, A. 1967. De la langue gu'on parle aux enfants roumains. In *To Honor Roman Jakobson*, Vol. 1.133-40.

Austerlitz, R. 1956. Gilyak nursery words. *Word* 12.260-79.

Bever, T. G., J. A. Fodor, and W. Weksel. 1965. Theoretical notes on the acquisition of syntax: A critique of 'contextual generalization.' *Psychological Review* 72.467-482.

Bhat, D. N. 1967. Lexical suppletion in baby talk. *Anthropological Linguistics* 9.33-6.

Blount, B. G. 1972. Parental speech and language acquisition: Some Luo and Samoan examples. *Anthropological Linguistics* 14.119-130.

Bolinger, D. 1964. Intonation as a universal. In H. Lunt (ed.), *Proceesings of the 9th International Congress of Linguists.* Cambridge, Massachusetts, 1962. The Hague: Mouton, 833-848.

Broen, P. 1972. The verbal environment of the language learning child. *Monograph of the American Speech and Hearing Association* 17.

Brown, R. and U. Bellugi. 1964. Three processes in the acquisition of syntax. *Harvard Educational Review* 34.133-151.

Brown, R., C. B. Cazden, and U. Bellugi. 1969. The child's grammar from I to III. In J. P. Hill (ed.), *Minnesota symposium on child psychology* Vol. 2.28-73. Minneapolis: University of Minnesota Press.

Byron, J. 1968. Berber nursery language. *Transactions of the Philological Society* 107-61.

Casagrande, J. B. 1948. Comanche baby language. *International Journal of American Linguistics* 14.11-14.

Cazden, C. B. 1965. Environmental assistance to the child's acquisition of grammar. Unpublished Ph.D. dissertation, Harvard University.

Chew, J. J. 1969. The structure of Japanese baby talk. *Journal-Newsletter of the Association of Teachers of Japanese* 6.4-17.

Chomsky, N. 1959. A review of *Verbal Behavior* by B. F. Skinner. *Language* 35.26-58.

Chomsky, N. 1967. The formal nature of language. Appendix A to E. H. Lenneberg, *Biological Foundations of Language.* New York: Wiley, 397-442.

Crawford, J. M. 1970. Cocopa baby talk. *International Journal of American Linguistics* 36.9-13.

Crystal, D. 1973. Non-segmental phonology in first language acquisition: Methodological problems. *Lingua* 32, No. 1-2.1-45.

Drach, K. 1969. The language of the parent: A pilot study. *Working Papers No.* 14, Language Behavior Research Laboratory, University of California, Berkeley.

Drachman, G. 1973. Baby talk in Greek. *OSU WPL* 15.

Ellis, J. and J. Ure. 1969. Language variety: register. In A. R. Metham, *Encyclopedia of Linguistics information and control.*

Ervin-Tripp, S. M. 1970. Discourse agreement: How children answer questions. In J. R. Hayes (ed.), *Cognition and the Development of Language*. New York: Wiley, 79-107.

Fairbanks, G. 1940. *Voice and Articulation Drillbook*. New York: Harper.

Farwell, C. 1973. The language spoken to children. *Papers and Reports on Child Language Development*. Stanford University 5.31-62.

Ferguson, C. A. 1956. Arabic baby talk. In Morris Halle, et al., (eds.), *For Roman Jakobson*. The Hague: Mouton 121-128.

Ferguson, C. A. 1964. Baby talk in six languages. *American Anthropologist* 66.103-114.

Ferguson, C. A. 1971. Absence of the copula and the notion of simplicity. In D. Hymes (ed.), *Pidginization and Creolization of Language*. Cambridge: Cambridge University Press, 141-150.

Ferguson, C. A. 1972. Towards a characterization of English foreigner talk. Paper presented at the Third International Congress of Applied Linguistics, Copenhagen, Denmark.

Ferguson, C. A. 1974. Baby talk as a simplified register. Paper presented at the Conference on Language Input and Acquisition, Boston, Massachusetts.

Ferguson, C. A., D. B. Peizer, and T. E. Weeks. 1973. Model-and-replica phonological grammar of a child's first words. *Lingua* 31.35-65.

Fischer, J. L. 1970. Linguistic socialization: Japan and the United States. In R. Hill and R. Konig (eds.), *Families in East and West*. The Hague: Mouton, 107-119.

Gleason, J. 1973. Code switching in children's language. In T. E. Moore (ed.), *Cognitive development and the acquisition of language*. New York: Academic Press, 159-168.

Grewel, F. 1959. How do children acquire the use of language? *Phonetica* 3.193-202.

Halliday, M. A. K., A. McIntosh, and P. D. Strevens. 1964. *The Linguistic Sciences and Language Teaching*. London: Longmans.

Henry. F. 1948. Discrimination of the duration of a sound. *Journal of Experimental Psychology* 38.734-743.

Henzl, V. 1973. Linguistic register of foreign language instruction. *Language Learning* 23.207-222.

Horii, Y. 1972. Fundamental frequency measures estimated from analysis of one sentence voice samples. *Asha* 14.466.

House, A. S. and G. Fairbanks. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America* 25.105-113.

Ivić, P. and I. Lehiste. 1969. Prilozi ispitivanju fonestke i fonološke prirode akcenata u savremenom srpskohrvatskom jeziku IV. *Zbornik za filologiju: lingvistiku* 12.115-165.

Kelkar, A. 1964. Marathi baby talk. *Word* 20.40-54.

Kobashigawa, B. 1969. Repetitions in a mother's speech to her child. *Working Papers* No. 14, Language Behavior Research Laboratory, University of California

Kones, R. ms. How to use the pitch extractor with the computer. Phonetics Laboratory, University of California, Berkeley.

Labov, W. 1970. The study of language in its social context. *Studium Generale* 23.30-87.

Lehiste, I. 1970. *Suprasegmentals*. Cambridge, Mass.: MIT Press.

Lehiste, I., and G. E. Peterson. 1961. Some basic considerations in the analysis of intonation. *Journal of the Acoustical Society of America* 33.419-425.

Linke, C. E. 1953. A study of pitch characteristics of female voices and their relationship to vocal effectiveness. Unpublished Ph.D. dissertation, University of Iowa.

Markel, N. N. 1965. The reliability of coding paralanguage: Pitch, Loudness, and tempo. *Journal of Verbal Learning and Verbal Behavior* 4.306-308.

McNeill, D. 1970. *The acquisition of language: The study of developmental psycholinguistics*. New York: Harper & Row.

Peterson, G. E. and H. L. Barney. 1952. Control methods used in a study of vowels. *Journal of the Acoustical Society of America* 24.175-184.

Pfuderer, C. 1969. Some suggestions for a syntactic characterization of baby talk style. *Working Papers* No. 14, Language Behavior Research Laboratory, University of California, Berkeley.

Phillips, J. R. 1970. Formal characteristics of speech which mothers use to address their young children. Unpublished Ph.D. dissertation, Johns Hopkins University.

Phillips, J. R. 1973. Syntax and vocabulary of mothers' speech to young children: Age and sex comparisons. *Child Development* 44.183-185.

Remick, H. L. 1971. The maternal environment of linguistic development. Unpublished Ph.D. dissertation, University of California at Davis.

Ruhm, H. B., E. O. Mencke, B. Milborn, W. A. Cooper, Jr., and D. E. Rose. 1966. Differential sensitivity to duration of acoustic stimuli. *Journal of Speech and Hearing Research* 9.371-384.

Ruke-Dravina, V. 1961. Ns. lastenhoitajain kielesta' ("On so-called nursery language"). *Virittaja* 1.85-91.

Sachs, J., R. Brown, and R. Salerno. 1972. Adults' speech to children. Paper presented at International Symposium on First Language Acquisition. Florence, Italy.

Skinner, B. V. 1957. *Verbal Behavior*. New York: Appleton-Century-Crofts.

Slobin, D. I. (ed.), 1967. *A field manual for cross-cultural study of the acquisition of communicative competence*. Berkeley: University of California, ASUC Bookstore.

Slobin, D. I. 1969. Questions of language development in cross-cultural perspective. *Working Papers* No. 14, Language Behavior Research Laboratory, University of California, Berkeley.

Snidecor, J. C. 1951. The pitch and duration characteristics of superior female speakers during oral reading. *Journal of Speech and Hearing Disorders* 16.44-52.

Snow, C. E. 1972. Mothers' speech to children learning language. *Child Development* 43.549-65.

Snow. C. E. 1974. Mothers' speech research: An overview. Papers presented at the Conference on Language Input and Acquisition. Boston, Massachusetts.

Stott, L. H. 1935. Time-order errors in the discrimination of short tonal durations. *Journal of Experimental Psychology* 18.741-766.

Voegelin, C. F., and F. M. Robinett. 1954. 'Mother language' in Hidatsa. *International Journal of American Linguistics* 20.65-70.

Vorster, J. 1974. Mothers' speech to children: Some methodological considerations. M.A. thesis, Institute for General Linguistics, University of Amsterdam

Waterson, N. 1971. Child phonology: A prosodic view. *Journal of Linguistics* 7.179-211.

Willis, D. D. 1974. Participant deixis in English and childish. Paper presented at the Conference on Language Input and Acquisition. Boston, Massachusetts.